THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Genome wide analysis and allele mining of diverse sorghum genes involved in nitrogen use efficiency.

Eugene Diatloff

B. Agr. Sc., Ph.D.

A thesis submitted for the degree of Master of Philosophy at

The University of Queensland in 2017

School of Agriculture and Food Science

ABSTRACT

Nitrogen (N) is the most important plant nutrient and is the most commonly applied element to agricultural crops. In many cases nitrogen fertiliser is the main input cost for farmers, and in addition, is a major pollutant from agricultural activity. Most plants only take up 50% or less of the applied N, and most of the remainder is lost to leaching into ground water and waterways, or volatilised into the atmosphere. Hence, improving the nitrogen responsiveness of crops is crucial for food security and environmental sustainability, and breeding N use efficient (NUE) crops has to exploit genetic variation for this complex trait.

In this thesis, reverse genetics was used to examine allelic variation in two key N metabolism genes. *In silico* analysis of the genomes of 44 genetically diverse sorghum genotypes identified a nitrate reductase and a glutamate synthase gene (NADH-GOGAT) that were under balancing selection in improved sorghum cultivars. It was hypothesised that these genes are a potential source of differences in NUE, and parents and progenies of nested association mapping populations were selected with different allelic combinations for these genes. Allelic variation was sourced from African (Macia) and Indian (ICSV754) genotypes that had been backcrossed into the Australian elite parent R931945-2-2. Nine genotypes with different allelic combinations were grown for 30 days in a glasshouse and supplied with continuous limiting (1 mM nitrate) or replete N (10 mM nitrate), or replete N for 27 days followed by three days N starvation prior to harvest. Biomass, N and nitrate contents were quantified together with gene expressions in leaves, stems and roots.

Limiting N supply universally resulted in less shoot and root growth, increased root weight ratio, reduced tissue nitrate and N concentrations, and reduced NADH-GOGAT expression. None of the tested genotypes exceeded growth or NUE of elite parent R931945-2-2. This may indicate that the allelic combinations did not confer an advantage during early vegetative growth, or that selection in a modern plant breeding program has already optimised the allelic combinations for these loci. It is also noteworthy that plants were grown under controlled environment conditions, and field responses may have been somewhat different. It is also possible that any selective advantage of other allelic combinations may only have been apparent in plants grown to anthesis and/or grain maturity. Thus, the next steps for ascertaining potential effects on NUE include growing plants to maturity. It is concluded that reverse genetics that take advantage of rapidly expanding genomic databases contributes towards a systematic approach for developing N efficient crops.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

No publications during candidature.

Publications included in this thesis

No publications included in this thesis.

Contributions by others to the thesis

Associate Professor Emma Mace conducted the population statistical analyses (Table 2.3) and this has been acknowledged in the thesis (section 2.2.1).

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

I would like to thank very much Professors Ian Godwin and Susanne Schmidt for allowing me to undertake this thesis and for their support, advice and encouragement throughout. This project was conducted under the University of Queensland Master of Philosophy post-graduate program.

My thanks to the following people:-

Assoc. Prof. Emma Mace and Dr. Shuaishuai Tai for bioinformatic assistance
Dr. Alan Cruickshank and Prof. David Jordan for providing the NAM population genotypes
Melissa Alves for assistance during harvest
David Appleton for LECO analyses
Belinda Worland and Dr. Nicole Robinson for helpful discussions.

A very special thanks to Sophie and my parents for all their support, and to little Lara for ensuring I did not waste too much time sleeping.

Keywords

allelic variation, Nested Association mapping population, nitrogen use efficiency (NUE), vegetative, nitrate reductase, glutamate synthase, sorghum

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 060705 Plant Physiology (60%)
ANZSRC code: 060405 Gene Expression (incl. Microarray and other genome-wide approaches) (20%)
ANZSRC code: 060411 Population, Ecological and Evolutionary Genetics (20%)

Fields of Research (FoR) Classification

FoR code: 0607 Plant Biology 60%
FoR code: 0604 Genetics 40%

TABLE OF CONTENTS Page

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations used in the thesis

N	Nitrogen
NUE	Nitrogen use efficiency
DW	dry weight
SA	surface area
SLN	specific leaf nitrogen
FW	fresh weight
NAM	Nested Association Mapping

INTRODUCTION

Nitrogen (N) is essential for plant biomass production, grain yield and quality. Nitrogen fertilizer together with plant genetics, were major factors that contributed to the "green revolution" which enabled the immense growth in the human population (Erisman et al. 2008). During this time, research on N fertilizer use mainly addressed questions about the "rates and dates" of fertilizer application, and the success of this approach cannot be denied. However, now a new set of questions are being asked of plant physiologists and breeders that address the efficiency with which N fertilizer is used. These questions have come about due to environmental concerns of the use and misuse of N fertilizer and the increasing cost of N fertilizer production. While there are farm management strategies available to improve agronomic nitrogen use efficiency (NUE) such as variable (spatial) and split (temporal)-fertiliser application, farmers do not yet have available to them plant genotypes which have enhanced NUE. Hence, there is an urgent need to provide such varieties.

Despite considerable research on NUE, there has been slow progress towards the release of improved NUE cultivars because of varying definitions of NUE (Sadras and Lemaire, 2014), the difficulty in screening for NUE, the considerable Genotype x Environment x Management interaction and the incomplete understanding of the molecular basis underlying NUE. Our understanding of the molecular basis of NUE has recently been advanced by the release of numerous genomes of both agricultural and model plant species. This genomic data can now be searched for genes involved in nitrogen acquisition, assimilation and metabolism, and their role in NUE assessed. In 2009, the first draft of the *Sorghum bicolor* genome was published (Paterson et al., 2009). More recently, the high coverage re-sequenced genomes of 44 sorghum lines representing the primary gene pool and spanning dimensions of geographic origin, end-use and taxonomic groups were presented (Mace et al., 2013). This provides an unmatched genomic resource covering the novel diversity available in sorghum enabling the detailed investigation of genes involved in NUE. This is particularly important since, although there have been some genetic improvements in sorghum, the amount of resources invested in the crop has been minimal compared to other cereal crop species (Dillon *et al.*, 2007). In addition, sorghum is an important crop in the semi-arid regions of the world and productivity is limited by soil fertility, especially N. Consequently, identification of new sources of genetic variability is essential to develop new cultivars with increased adaptation to abiotic stresses such as low soil N.

Numerous studies have examined the nitrogen responses of a wide range of sorghum genotypes (eg. Maranville *et al.*, 2002a, b; Youngquist *et al.*, 1992). An alternative approach, which has not yet been reported in sorghum, is to phenotype genotypes that contain allelic variation for genes specifically involved N uptake, assimilation and metabolism in sorghum. Selection of such genotypes is now enabled by the availability of genetic resources such as the sorghum nested association mapping (NAM) populations created by Jordan et al. (2011), and the availability of genomic sequencing data by Mace et al. (2013).

.

This project consists of two parts:-

1) *In silico* - the *in silico* analysis of the whole genome sequence data from 44 diverse sorghum genotypes to identify genes involved in NUE that are under selection, and

2) Experimental - the subsequent phenotyping of sorghum genotypes that contain contrasting alleles for these genes of interest.

CHAPTER 1
LITERATURE REVIEW

1.0 Sorghum

1.1.1 Introduction

Grass species are important worldwide as staple grain foods for humans, feedstock for animals and as a source of biomasss. Sorghum (*Sorghum bicolor* (L.) Moench) is the fifth most important cereal crop globally (after wheat, rice, maize and barley, FAO 2012) that can survive the harsh climatic conditions of arid and tropical environments. Unlike maize and other staple crops, which are widely used for both food and industrial purposes, approximately 90% of the world's sorghum growing area lies in Africa and Asia (FAO 2012) where it has remained mainly a traditional food and multipurpose crop of subsistence farmers. Genetic improvement of sorghum has lagged behind other cereal crops (Dillon *et al.*, 2007) and consequently presents an immense potential opportunity for plant breeders and plant physiologists.

1.1.2 Taxonomy and comparative phylogeny

The grasses are angiosperms, and are members of the large monocot clade, which includes about 20% of known flowering plants. Within the monocots, the grasses belong to the commelinid clade (Fig. 1.1).

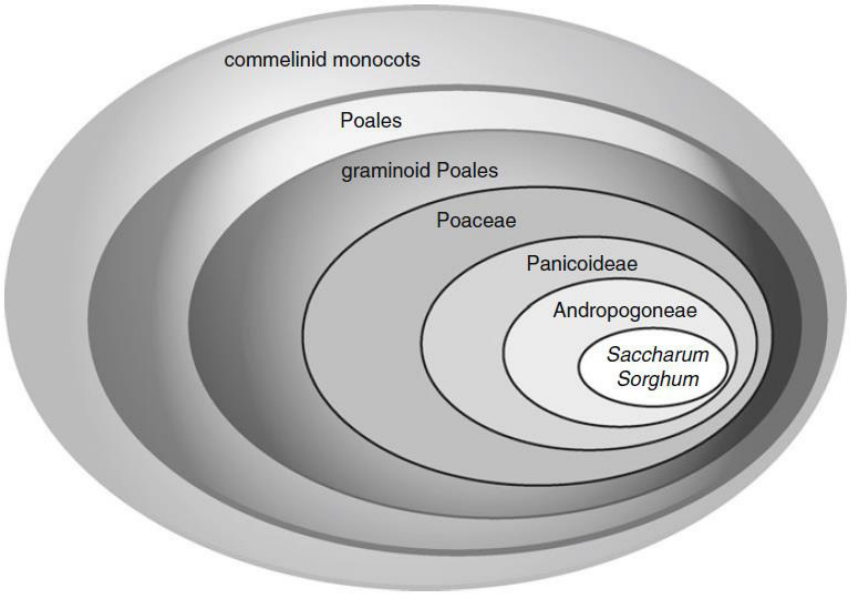

Figure 1.1 Venn diagram showing the nested relationships of the commelinid clade, Poales, and included taxa (Kellogg 2013).

Sorghum (along with maize, sugarcane and miscanthus) belongs to the Andropogoneae (Fig. 1.1), whose members all have the C_4 photosynthetic pathway NADP-ME (NADP-malic enzyme), with single bundle sheaths (see section 1.2.4 Nitrogen Use Efficiency, for discussion on C_4 plants and NUE). Andropogoneae include ca. 1000 species, representing approximately one-tenth of the world's grass species and are most diverse in the Old World (Skendzic 2013). The poor phylogenetic resolution of Andropogoneae, and general lack of appropriate taxon sampling, means that we do not know precisely where *Saccharum*, *Miscanthus* and *Sorghum* fall within the tribe and what their closest relatives are (Kellogg 2013).

Cultivated sorghum is classified into five main races (bicolor, guinea, caudatum, durra and kafir) and 10 intermediate races based on pair-wise combinations of the five main races (Barnaud *et al.* 2008, Harlan and Dewet 1972). The identification of racial divisions and species are primarily based on panicle and grain characters (Harlan and Dewet 1972). Sorghum panicles, which are called inflorescences, show remarkable diversity in morphological, physiological, genetic and ecological traits. Sorghum inflorescence architecture is not only an important factor for sorghum identification, but contributes to both the yield and quality of sorghum.

DNA sequencing has now allowed comparative genomics and phylogeny. The ancestral sorghum genome and the maize progenitor genomes divergence is approximately 12 million years ago, while sugarcane divergence from sorghum is about 5 million years ago (Figure 1.2; Muraya 2014). Gene orders appear to be largely conserved between sorghum and maize; only a limited number of rearrangements have been identified (Singh and Lohithaswa 2006). Sugarcane and sorghum appear to be more closely related than either is to maize (Singh and Lohithaswa 2006).

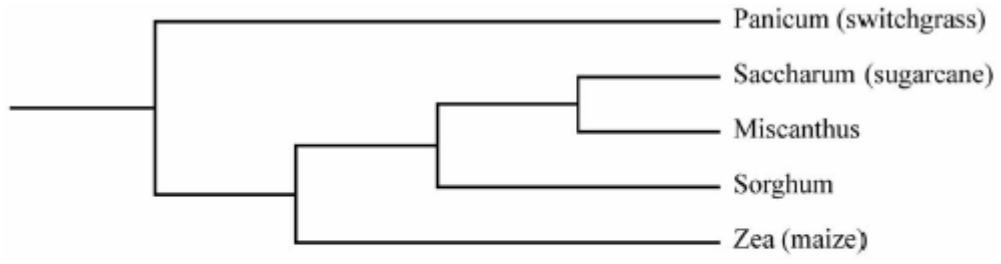

Figure 1.2 Phylogenetic tree depicting the relationship between the C_4 grasses maize, miscanthus, sorghum, sugarcane and switchgrass (Adopted from Lawrence and Walbot (2007); no time scale presented by authors).

1.1.3 Distribution

Ethiopia and Sudan are believed to be the centre of origin and domestication of sorghum, and hence sorghum has been referred to as Africa's indigenous cereal crop. Roughly 90% of the world's sorghum growing area lies in developing countries mainly in Africa and Asia (FAO 2012). Sorghum has spread over the drier areas of the world and is now even being considered as an alternative to maize in Europe for biofuel production. The predicted future climate warming and reduced rainfall in some areas would suggest that sorghum may play an even more important role in many countries.

1.1.4 Genetics

Sorghum is a self-pollinating diploid (2n = 2x = 20) with a genome about 25% the size of maize and sugarcane (see section 1.1.5 Genome). Though there have been some improvements in sorghum, the amount of resources invested in the crop has been minimal compared to other species. In addition, it has been identified in Australia that a genetic bottleneck has been created in cultivars due to the consistent breeding for sorghum midge resistance (Jordan *et al.* 1998). Consequently, there is an urgent need to introduce new genetic material.

1.1.5 Genome

The genome of the leading US hybrid parents, BTx623 was published in 2009 (Paterson *et al.* 2009). The genome is approximately 732 Mb in size (≈47,000 protein coding transcripts) (https://phytozome.jgi.doe.gov). Subsequent to the first publication of the sorghum genome, there have been other publications of other sorghum genomes including the recent report on the resequencing of 44 diverse sorghum genotypes (Mace *et al.* 2013). The genome sequence data of these 44 sorghum genotypes provides a rich source of data to be mined for allele diversity. This genomic data combined with NAM populations (Nested Association Mapping) based on some of the 44 sequenced genotypes (Jordan *et al.* 2011) provide an excellent resource to investigate and dissect the genetics and molecular biology of biotic and abiotic stresses and agronomically important traits such as nitrogen use efficiency.

1.1 Nitrogen

1.2.1 Nitrogen cycle

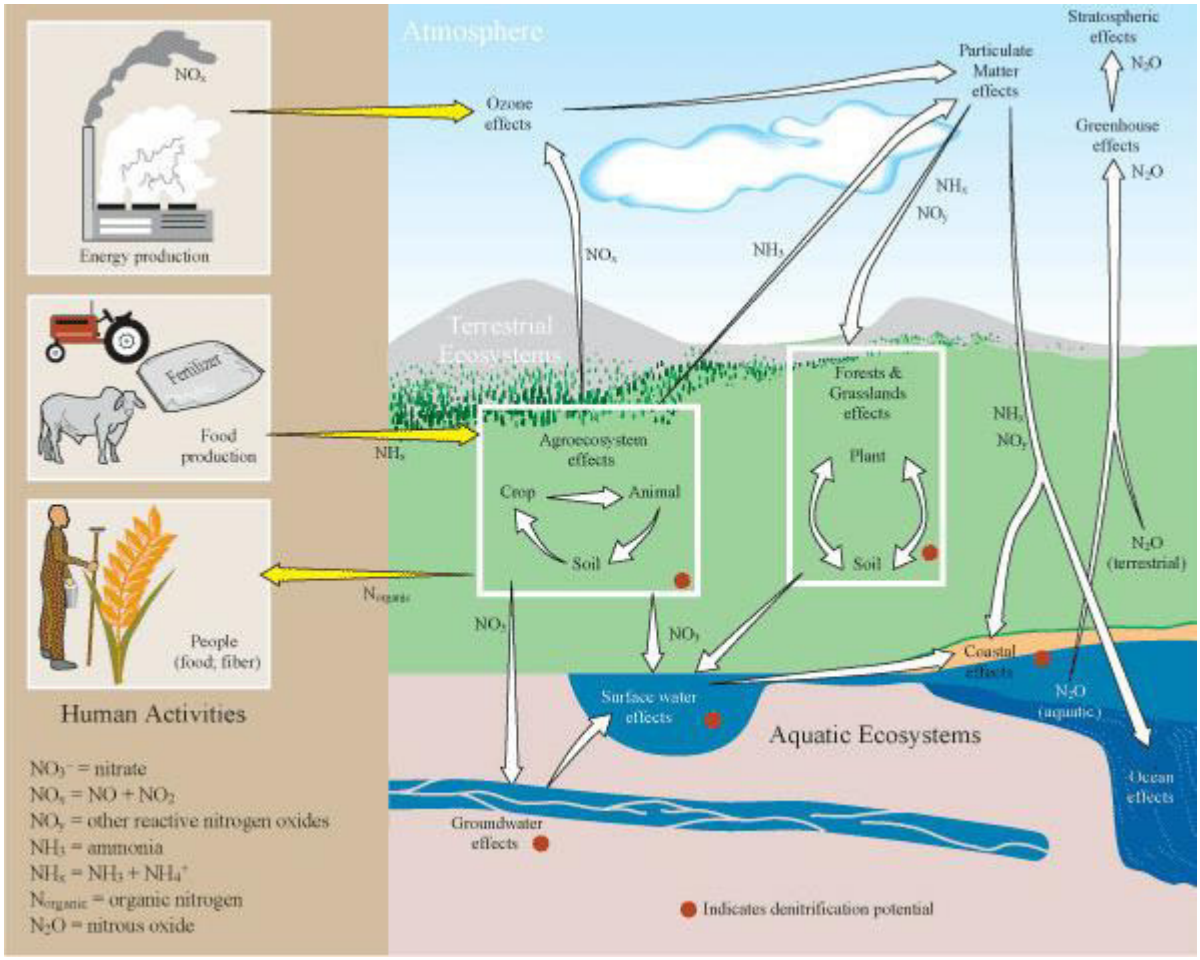

Figure 1.3 Example of the nitrogen cycle. (http://www.unep.org/yearbook/2003/ilust57.htm)

Figure 1.3 gives an overview of the nitrogen cycle. Although the pathways and cascades have been studied in some detail, the quantification of the flows and transformations are difficult and the values very uncertain (Gruber and Galloway 2008). The problems of pollution (soil, water, air) by reactive nitrogen compounds has put more focus on the use and misuse of nitrogen fertilizers and consequently the nitrogen nutrition of plants.

1.2.2 Nitrogen nutrition of sorghum

Warncke and Barber (1974) found that sorghum and maize were capable of absorbing nitrate down to similar very low concentrations in solution culture. Thus sorghum roots reduced the nitrate concentration to 2.7 µM before uptake ceased whereas the range for three maize cultivars was 2 to 4 µM. However, Forno (1977) found that sorghum required higher nitrate concentrations

in the root environment for maximum growth. When N was supplied as ammonium, the external concentration required for maximum vegetative growth by sorghum was again higher than for maize.

Grain N concentration is an important quality factor in sorghum. In precisely controlled solution culture experiments, N supply had major effects on grain number, yield and on grain N concentration (Cowie 1973; Asher and Cowie 1974). Plants subjected to N deficiency during the vegetative phase between planting and floral initiation produced only a small panicle with fewer primary branches, secondary branches, and visible florets at head emergence than control plants supplied with adequate N. During the reproductive phase, N stress between floral initiation and anthesis caused between 16 and 30% of the initiated florets to abort. Nitrogen stress following anthesis had little effect on grain yield but greatly reduced grain N concentration compared with plants receiving adequate N. With continuous N stress, the reductions in grain number due to reduced floral initiation and subsequent abortions brought the grain number into sufficient balance with the N supply to produce grain of acceptable N content.

1.2.3 Nitrogen metabolism

Nitrogen management in plants can be divided based on the two differing growth stages, vegetative and reproductive (Hirel *et al.* 2007). During the vegetative phase, nitrate is absorbed, assimilated into amino acids then proteins that are used for growth and metabolism. Upon switching to the reproductive phase, proteins are degraded into amino acids and ammonium and transported to the developing reproductive organs and stored.

Most of our knowledge on regulation of N metabolism has been derived from C_3 plants and no differences in organization and regulation of nitrate assimilation between C_3 and C_4 plants were described (Kopriva 2011). An example of the N pathway and gene families found in cereals is presented in Figure 1.4.

C_4 plants differ significantly from C_3 plants in the compartmentalization of N metabolism and also in N use efficiency. It has long been known that in terms of growth rate C_4 grasses respond better to applied N than C_3 grasses (Hallock *et al.* 1965). The C_4 species clearly exhibited higher N use efficiency, expressed as biomass per unit N in plant (Brown 1978). The greater N use efficiency seems to be connected with a lower content of Rubisco in the leaves due to the CO_2 concentrating mechanism (see section 1.2.4 below).

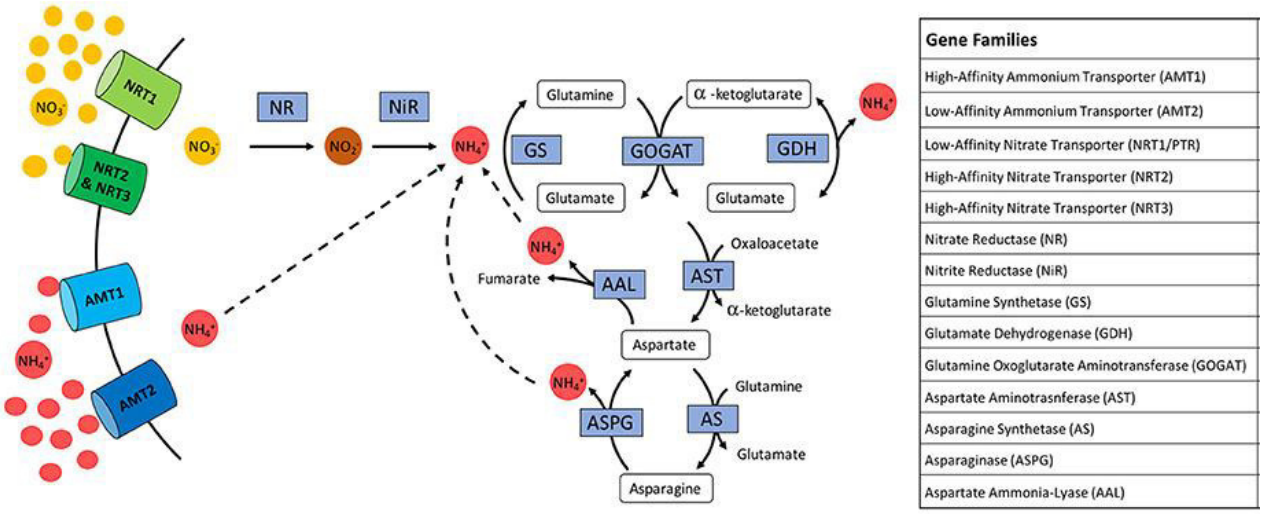

Figure 1.4 Diagram of biosynthetic pathway in cereals of N uptake and utilization to generate the amino acids glutamine, glutamate, aspartate, and asparagine (Massel et al. 2016)**.**

It is interesting to note that rising CO_2 concentrations cause inhibited nitrate assimilation in wheat plants (Bloom *et al.* 2014), reducing their productivity and protein levels. However, Bloom et al. (2012) also showed that CO_2 enrichment inhibits shoot nitrate assimilation in C_3 plants (including wheat, barley Arabidopsis), but not C_4 plants (including maize), and slows growth under nitrate in C_3 plants. Consequently, growing crop varieties with improved NUE to counteract rising CO_2 concentrations may only apply to C_3 plants, although the effect of CO_2 increases on nitrate assimilation in sorghum does not appear to have been investigated.

1.2.4 Nitrogen use efficiency (NUE)

Sorghum is a representative C_4 plant harbouring plastidic NADP-ME (NADP-malic enzyme) for C_4 photosynthesis (as is maize); its capacity to assimilate and metabolize C and N compounds is greater than that of C_3 plants (Ghannoum *et al.* 2011, Kiirats *et al.* 2010). Because C_4 species concentrate CO_2 at the site of Rubisco, the theoretical requirement for nitrogen in photosynthesis is less than in C_3 species. The benefit of a lower requirement for Rubisco in C_4 leaves is, however, partially offset by the N requirement for the enzymes of C_4 metabolic cycle, primarily PEP carboxylase and PPDK (Sage *et al.* 1987). At the whole plant level the difference in leaf nitrogen concentration between C_3 and C_4 plants combined with the higher leaf photosynthetic rate of C_4 species results in a photosynthetic NUE that is approximately twice as high in C_4 compared to C_3 plants (Brown 1978, Pearcy and Ehleringer 1984). Direct comparisons between sorghum and maize in their NUE have resulted in differing findings. Muchow (1988) did not find

any difference in NUE components between maize and sorghum in field trials in Queensland. However, in France a comparison of N uptake capacities of maize and sorghum under contrasting soil N availability showed that under non-limiting N supply, the two crops have similar N uptake, while under severe N limitation the N uptake capacity of sorghum is higher than that of maize (Lemaire *et al.* 1996). The comparisons by Muchow (1988) were carried out at equivalent growth duration, whereas those of Lemaire et al. (1996) were at equivalent biomass. The reasons for the higher N uptake capacity by sorghum under severe N limitation observed by Lemaire et al. (1996) is unclear, but Hirel et al. (2007) suggested that it could be due to a more developed and branched root system for sorghum as compared with maize.

Assessing how effectively crops take up and use nitrogen is complex (Fischer *et al.* 2014). The definition of NUE is very much dependent on the scale being described, ranging from the cellular to the ecosystem level and including the tissue, organ, plant, and plant community levels/scales. Good et al. (2004) provide a list of some of the definitions and formulae used to describe nitrogen use efficiency (Table 1.1).

Table 1.1 Definitions and formulae used to describe nutrient use efficiency in plants (Good et al., 2004).

Eqn	Term	Formula	Definition	Comments
1	Nitrogen use efficiency	$NUE = Sw \div N$	Sw, shoot weight (DW); N, nitrogen content of shoots (DW)	Does not account for biomass increases
2	Usage index	$UI = Sw \times (Sw \div N)$	Sw, shoot weight; N, nitrogen in shoots	Takes into account absolute biomass increase
3	Nitrogen use efficiency (grain)	$NUE = Gw \div Ns$	Gw, grain weight; Ns, nitrogen supply (g per plant)	Reflects increased yield per unit applied nitrogen
4	Uptake efficiency	$UpE = Nt \div Ns$	Nt, total nitrogen in plant; Ns, nitrogen supply (g per plant)	Measures efficiency of uptake of nitrogen into plant
5	Utilization efficiency	$UtE = Gw \div Nt$	Gw, grain weight; Nt, total nitrogen in plant	Fraction of nitrogen converted to grain
6	Agronomic efficiency	$AE = (Gw_F - Gw_C) \div N_F$	N_F, nitrogen fertilizer applied; Gw_F, grain weight with fertilizer; Gw_C, grain weight of unfertilized control	Measures the efficiency of converting applied nitrogen to grain yield
7	Apparent nitrogen recovery	$AR = (N_F \text{ uptake} - N_C \text{ uptake}) \div N_F \times 100$	N_F uptake = plant nitrogen (fertilizer); N_C uptake = plant nitrogen (no fertilizer); N_F = Nitrogen fertilizer applied	Measures the efficiency of capture of nitrogen from soil
8	Physiological efficiency	$PE = (Gw_F - Gw_C) \div (N_F \text{ uptake} - N_C \text{ uptake})$	Gw_F, grain weight (fertilizer); Gw_C, grain weight (no fertilizer)	Measures the efficiency of capture of plant nitrogen in grain yield

Significant genotypic differences for NUE have been documented in sorghum (Table 1.2).

Table 1.2 Table of studies examining NUE of sorghum.

Equation (Table 1.1)	Number of genotypes	Nitrogen rates	Reference
1, 3	12	116 kg N/ha	(Maranville *et al.* 1980)
1	4	Low and high soil N	(Maranville and Madhavan 2002)
1, 3, 6 & 8	3	0, 45, 90, 135 & 180 kg N/ha	(Maranville *et al.* 2002)
1	5	Low and medium soil N	(Gardner *et al.* 1994)
1, 3 & 6	5	0, 60, 120, 180 & 240 kg N/ha	(Hibberd and Hall 1990)
1, 3	14	0, 60 & 240 kg N/ha	(Kamoshita *et al.* 1998a)
3	4	0 & 60 kg N/ha	(Kamoshita *et al.* 1998b)
3	3-6	0, 60 & 240 kg N/ha	(Kamoshita *et al.* 1998c)

For example, Gardner et al (1994) showed that the improved pure line bred in India (M35-1) was the most nitrogen efficient because it was able to partition more dry matter into stalk tissue, maintain thicker leaves, and most rapidly remobilized N from older to younger leaves. However, altogether these studies (Table 1.2) used only a total of 53 different genotypes which is surprisingly small compared to the number of genotypes screened for NUE in other crops such as maize. Even so, there is good reason to believe that improvements in NUE in sorghum can be achieved using genetic approaches. The availability of the sorghum genome sequence (Paterson et al. 2009), facilitates genotyping the mapping populations using whole genome sequencing approaches in search of genes associated with NUE.

1.2.5 Genes associated with NUE

Genes associated with NUE in crop and model plants have been identified using a number of techniques including both forward genetic (phenotype to genotype, e.g. QTLs (Quantitative Trait Loci) and QTL mapping) and reverse genetic (genotype to phenotype, e.g. gene knock-outs) approaches. For sorghum, only the forward genetic approach using QTLs has been used to identify genes involved in NUE. Gelli et al. (2014, 2016) used a combination of QTLs and transcriptional profiling to search for genes involved in NUE traits in 2 RIL populations (CK60 “low- NUE US line” x China17 “higher NUE Chinese line” and Ck60 x San Chi San “low-N tolerant Chinese line”)

evaluated under low (0 kg N.ha^{-1}) and normal N (100 kg N. ha^{-1}) conditions. The RNA-seq data showed that one QTL for NUE contained differentially expressed gene transcripts encoding ferredoxin-nitrite reductase (FNR), chloroplast localized serine/threonine-protein kinase (SNT7), and a SufE/NifU family protein. FNR and SNT7 were highly expressed in China17 (higher NUE Chinese line) than in CK60 (low- NUE US line). Gelli et al. (2014) showed that the region of chromosome 9 harbors the highly expressed gene encoding NADH-GOGAT and a glutamine-rich protein. However, these genes were not differentially expressed between CK60 (low- NUE US line) and China17 (higher NUE Chinese line) according to RNA-seq data. Recently, Gelli et al. (2016) evaluated QTLs and gene expression of 131 recombinant inbred lines (CK60 (low- NUE US line) x China17 (higher NUE Chinese line)) under normal and low N conditions. Co-localized regions affecting multiple agronomic traits were detected on chromosomes 1, 5, 6, 7 and 9. These potentially pleiotropic regions were coincident with the genomic regions of cloned QTLs, including genes associated with flowering time and plant height. In these regions, RNA sequencing data showed differential expression of transcripts related to nitrogen metabolism (ferredoxin-nitrate reductase), glycolysis (phosphofructo-2-kinase), seed storage proteins, plant hormone metabolism and membrane transport.

Over-expression of NADH-GOGAT in rice resulted in an increase in grain weight, indicating that NADH-GOGAT is indeed a key enzyme in nitrogen utilization and grain filling in rice (Yamaya et al. 2002). In wheat, Quraishi et al. (2011) validated the NUE QTL on chromosome-3B, and proposed that a GOGAT gene is conserved structurally and functionally at orthologous positions in rice, sorghum and maize genomes and that this gene likely contributes significantly to NUE in wheat and other cereals. It will be of interest to determine if breeding that allows for higher expression of nitrate assimilation and GOGAT can increase biomass and grain yield by increasing nitrate assimilation and amino acid production and incorporation into proteins.

It is important to emphasize that although forward genetics (phenotype to genotype) have been used in sorghum for NUE analysis, the use of reverse genetics (genotype to phenotype) such as the use of NAM populations have not been used to date.

The objectives of the following two chapters were:-

Chapter 2 - identification of genes involved in NUE that are under selection in a panel of 44 diverse sorghum genotypes and selection of NAM parents and progeny that contain different combinations of contrasting alleles for those genes,

Chapter 3 - phenotyping the sorghum lines selected in Chapter 2 for their responses to varying root zone nitrate concentrations.

CHAPTER 2

SELECTION HISTORY OF GENES ASSOCIATED WITH NITROGEN UPTAKE, ASSIMILATION AND METABOLISM IN SORGHUM

2.1 Introduction

Developments in next generation sequencing technology have resulted in an unprecedented release of plant genome data. Initially this involved only the smaller genomes of model plant species, but now includes many important agricultural, horticultural, forage, wood producing and pharmaceutical plants. All these data, combined with user friendly bioinformatic tools has enabled the easy access, search and analysis of plant genome data.

The first release of the *Sorghum bicolor* genome (Paterson et al 2009) was followed by the release of the genomes of a number of sorghum genotypes (Mace et al. 2013; Luo et al. 2016; Bekele at el., 2013). Mace et al. (2013) presented genomic data from parental lines of a recently developed sorghum Nested Association Mapping population (Jordan et al., 2011). Together these resources will facilitate the dissection of complex traits such as NUE and the identification and exploitation of SNPs associated with favourable variants. The following chapter reports the use of *in silico* analysis to identify those genes associated with NUE that are also under selection in the population of 44 sorghum genotypes re-sequenced by Mace et al. (2013).

2.2 Materials and methods

Mace et al. (2013) selected 44 accessions to represent all major races of cultivated *S. bicolor*, in addition to its progenitors and *S. propinquum*. Eighteen lines were considered to be landraces, 17 were improved inbreds, seven were wild and weedy sorghums, and two were *S. propinquum*.

The genomes of these 44 sorghum genotypes have been deposited in the NCBI Short Read Archive under accession codes SRS378430 to SRS378473. The extraction of nucleotide sequence data was performed via requests through Assoc. Prof. Emma Mace (QAAFI) and Dr. Shuaishuai Tai (BGI-Shenzhen). The supplied genomic data was analysed using the programs presented in Table 2.1.

Table 2.1 Computer programs used to analyse DNA and protein sequence data.

Analyses conducted	Computer program/suite	Where analyses were conducted	Reference
Gene and protein prediction	AUGUSTUS	online[1]	Stanke & Morgenstern 2005
Protein annotation: functional categories	Blast2GO	online[2]	Gotz et al. 2008
Multiple sequence alignment	MUSCLE	online[3]	Dereeper et al. 2008
Phylogeny	PhML	online[3]	Dereeper et al. 2008
Tree viewers	TreeDyn EvolView	online[3] online[4]	Dereeper et al. 2008 He et al 2016

[1] http://bioinf.uni-greifswald.de/webaugustus/

[2] https://www.blast2go.com/

[3] www.phylogeny.fr/

[4] http://www.evolgenius.info/evolview/#login

The aim of the *in silico* analysis part of this thesis was to identify NUE related genes that are under selection in any of the three groups of genotypes (improved inbreds; landraces; wild and weedy) in the panel of 44 diverse sorghum lines that have been sequenced (Mace et al., 2013).

Two strategies were employed to search for these genes of interest:-

a) NUE homologue genes under selection – first, taking the genes that have been identified in the scientific literature to be involved in NUE in other plant species (eg. maize, rice, *Arabidopsis*), and then testing if their sorghum homologues are under selection,

b) Genes under selection that are potentially NUE related – first, taking all the genes that have been identified from the sorghum whole genome sequencing data to be under selection, and then searching to see if any of these genes under selection are potentially involved in NUE.

2.2.1 Strategy a) NUE homologue genes under selection

An initial list of 149 genes were identified to be involved in NUE based on experimental evidence from other plant species presented in the scientific literature (Table 2.2). The nucleotide sequences from the 44 sorghum genotypes for all the 149 genes listed in Table 2.2 were subjected to statistical analyses to identify those under selection in any of the three groups of genotypes (improved inbreds; landraces; wild and weedy). Analysis was performed by Assoc. Prof. Emma Mace by evaluating nucleotide diversity ($\theta\pi$; Nei, 1987), Watterson's estimator ($\theta\omega$; Watterson, 1975), and neutrality test *Tajima's D* (TajD; Tajima, 1989) using BioPerl modules with an in-house script (Mace et al., 2013).

2.2.2 Strategy b) Genes under selection potentially NUE related

For this strategy, a list of 1213 genes that show domestication improvement features under selection are presented by Mace et al. (2013; Supplementary Data 10). In Supplementary Data 10 (List of candidate domestication and improvement features under selection) there are the following pair-wise comparisons:-

- Improved vs Landraces – 209 predicted gene models
- Improved vs Wild & Weedy – 418 predicted gene models
- Landrace vs Wild &Weedy – 586 predicted gene models

Total = 1,213

To determine which of these 1,213 genes are likely to be involved in NUE, a search was conducted (using the predicted amino acid sequences) to extract their putative functional annotations from the NCBI database (http://www.ncbi.nlm.nih.gov/). This search was conducted using the Blast2GO program (http://www.blast2go.com/b2ghome, Table 2.1). No additional genes to those listed in Table 2.2 directly involved in nitrogen metabolism were identified in this list of 1,213 genes.

Table 2.2 Number of sorghum homologues of NUE related genes identified from the literature and subjected to *in silico* analysis for evidence of selection in 44 diverse sorghum genotypes.

Category	Annotation	No. of sorghum genes identified
Transporters & channels		
	PTR/NRT family = (Peptide TRansporter/Nitrate TRansporter 1)	91
	Nitrate transporters (family NRT2 and NRT3)	7
	Ammonium transporters	8
	CLCs (ChLoride Channel)	10
Nitrogen assimilation enzymes		
	Nitrate reductase	2
	Nitrite reductase	1
	Glutamine synthetase	5
	Glutamate dehydrogenase	2
	Alanine amino transferase	2
	Aspartate aminotransferase	2
	Asparagine synthetase	1
	△1-pyrroline-5-carboxylate synthetase	1
	Mitochondrial folylpolyglutamate synthetase	1
	GoGAT - Glutamine oxoglutarate aminotransferase = glutamate synthase	3
	ATG – autophagy	1
Transcription factors		
	DRO1	3
	DOF1	2
	NLP	5
Nitrogen regulators		
	CIPK23	1
	CIPK8	1
	TOTAL =	149

2.2.3 Phylogenetic analyses

The nucleotide sequences genes considered to be under selection were subjected to phylogenetic analyses to visualize groupings of related sequences. Those genes that showed phylogenetic relationships with clearly delineated groupings amongst the genotypes and significant diversity, were further subjected to phylogenetic analyses of their predicted amino acid sequences to examine if the groupings observed at the DNA level persisted at the protein level. Nonsynonymous single nucleotide polymorphisms (nsSNPs) are coding variants that introduce amino acid changes in their corresponding proteins. These nsSNPs can affect protein function and hence the interest in comparing the phylogenetic patterns amongst the genotypes at both DNA and amino acid levels. The predicted amino acid sequences for gene DNA sequences from each of the 44 genomes were obtained using AUGUSTUS (Table 2.1) trained to the corresponding translated genes from the published genome of BTx623 (Paterson et al. 2009, https://phytozome.jgi.doe.gov/pz/portal.html). Phylogenetic analyses were conducted using http://www.phylogeny.fr/index.cgi (Table 2.1) using the preset parameters. Dendogram annotations were added using EvolView (Table 2.1).

2.3 Results and Discussion

From the *in silico* analyses, five genes were identified that, were associated with NUE, showed balancing selection in one or more of the three groups of genotypes (improved inbreds; landraces; wild and weedy) and similar groupings in the phylogenetic analyses of both nucleotide and predicted amino acid sequences (Table 2.3).

Two of the genes listed in Table 2.3 are involved in nitrogen assimilation, nitrate reductase (Sobic.004G196101) and NADH-GOGAT (Sobic.009G225700). The other three genes belong to the NPF (NRT/PTR Family = Nitrate transporter 1/Peptide transporter; Leran et al., 2014). Two of these genes from the NPF are located consecutively on Chromosome 6 (Sobic.006G130400 and Sobic.006G130501). However, there are insufficient functional data available in the scientific literature (neither on the genes nor on their homologues in other plant species), to determine which type of nitrogenous substrates (including glucosinolates) these three NPF proteins transport. This makes it difficult to design appropriate screens for phenotyping.

Table 2.3 The summary population statistics of sorghum NUE homologues showing selection and similar phylogenetic analysis patterns in both nucleotide and predicted amino acid sequences. Values highlighted with an asterisk (*) are greater than 2.

Gene	Annotation	Selection	Selection criteria								
			Tajima's D			$\theta\pi$ ($x10^{-3}$)			$\theta\omega$ ($x10^{-3}$)		
			Landraces	W & W	Improved	Landraces	W & W	Improved	Landraces	W & W	Improved
Sobic.004G196101	Nitrate reductase	Balancing in Improved	1.3	0.8	2.1*	2.6	3.7	3.3	1.7	2.5	1.8
Sobic.009G225700	NADH-GOGAT	Balancing in Improved & Landraces	3.0*	-1.2	2.2*	3.4	2.5	3.5	2.0	2.5	1.9
Sobic.006G130400	NPF4.13	Balancing in Landraces	2.6*	1.1	1.1	5.9	6.5	4.5	3.0	4.2	3.0
Sobic.006G130501	NPF4.12	Balancing in Landraces	2.2*	1.1	1.2	4.7	5.4	3.6	2.5	3.5	2.3
Sobic.010G133100	NPF7.8	Balancing in Improved	1.9	0.3	2.1*	6.4	5.5	6.4	3.6	4.0	3.6

Balancing selection was detected in the nucleotide sequences of the nitrate reductase (Sobic.004G196101) and NADH-GOGAT (Sobic.009G225700) genes (Table 2.3). These genes have also been identified to be under balancing selection by Massel et al. (2016). Balancing selection indicates the long-term selective maintenance of multiple alleles (Wright and Gaut, 2004; Delph and Kelly, 2014). In contrast to strictly advantageous or deleterious mutations, whose persistence times as polymorphisms are generally short, balanced polymorphisms can be maintained indefinitely. They are also more likely to be segregating at intermediate frequencies, where they contribute most to population variance affecting fitness. Thus, there are good reasons to be interested in identifying balanced polymorphisms in a species (Tian et al. 2002).

2.3.1 Sobic.004G196100 – Nitrate reductase (NR)

(gDNA = 3,104 bp with 3 introns; CDS = 2,772 bp; protein = 924 amino acids)

There are three genes in sorghum that are homologous to nitrate reductase (Figure 2.1). The sorghum NR homologues are Sobic.004G196100, Sobic.004G312500 and Sobic.007G153900 (Figure 2.1), with only the first being under selection, balancing selection in the Improved group (Table 2.3). The only NR genes shown in Figure 2.1 that have been functionally characterized using a combination of mutants and enzyme assays are the two Arabidopsis NR genes (NIA1_ At1g77760 and NIA2_At1g37130).

To examine the gene expression patterns of the Sorghum NR homologues, the MOROKOSHI sorghum transcriptome database was queried (http://sorghum.riken.jp/morokoshi/Home.html; Makita et al. 2015) and results summarized in Figure 2.2. The expression of Sobic.004G312500 is highest in roots compared to all other tissues analysed. The expression of Sobic.004G196100 is also highest in roots, but also in stems and the growing leaf sheath under certain conditions. The expression of Sobic.007G153900 is higher in roots and leaves than in other tissues. In general, the relative expression levels of the three genes is approximately; Sobic.004G312500≈Sobic.007G153900>>Sobic.004G196100.

The enzymatic assay for nitrate reductase activity requires the addition of nitrate to a protein extract and measuring the nitrite produced (Wray and Fido, 1990), however, the differentiation between different NR proteins within a specific plant tissue is not possible by enzymatic assay.

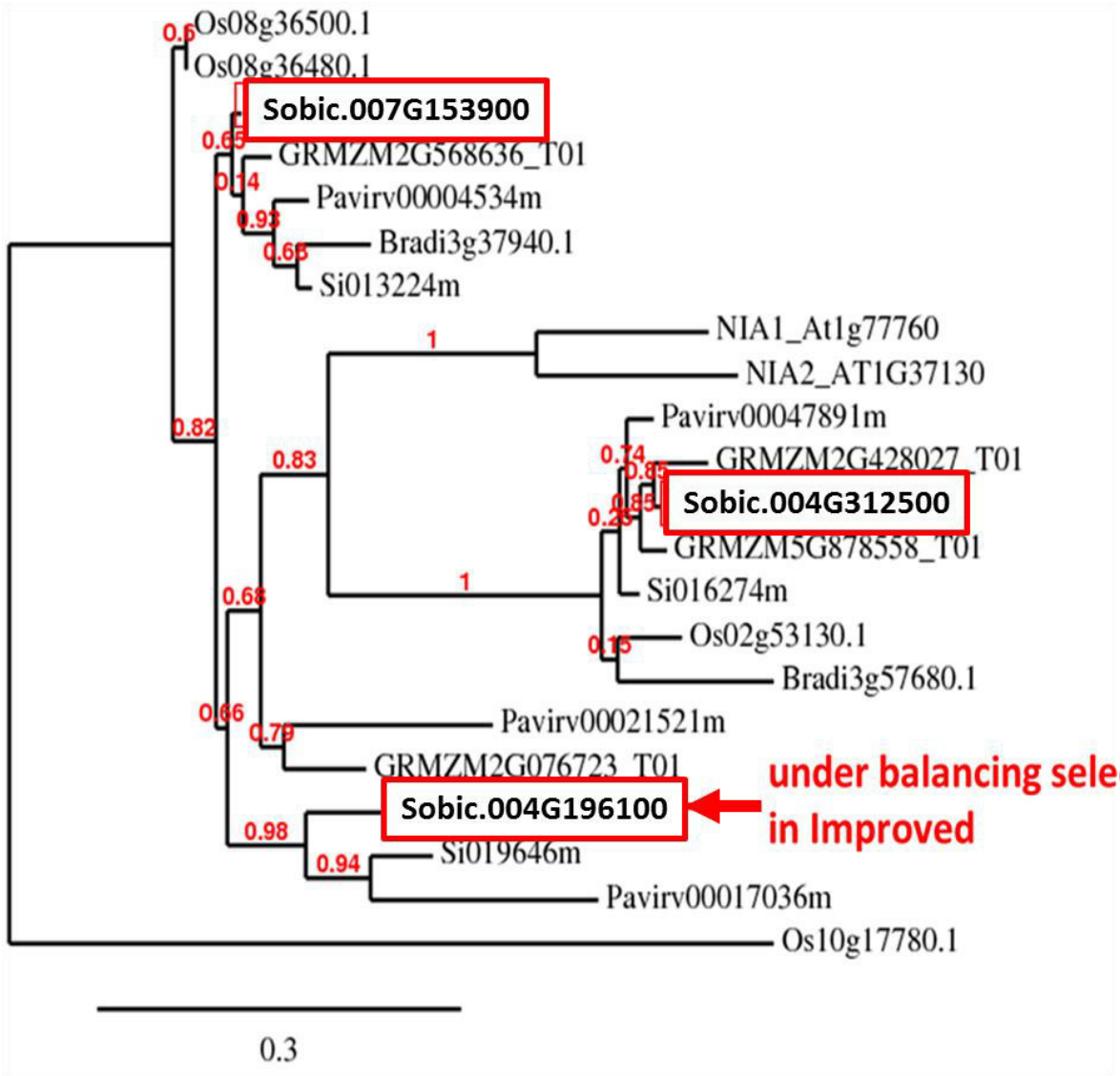

Figure 2.1 Phylogenetic relationship between the protein sequences of nitrate reductase in *Arabidopsis* and their homologues in the monocots with completed whole genome sequences.

(Protein sequences extracted from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) and phylogenetic analyses conducted at http://www.phylogeny.fr (Table 2.1).

Pavirv = *Panicum virgatum* Si = *Setaria italica* Bradi = *Brachypodium distachyon*

GRMZM = *Zea mays* Os = *Oryza sativa* At = *Arabidopsis thaliana*

Sobic = *Sorghum bicolor* Red box = *Sorghum bicolor* NR genes

Scale bar = 3 amino acid substitutions per 10 amino acids.

Red numbers represent a measure of support for the node; 0-1 where 1 represents maximal support.

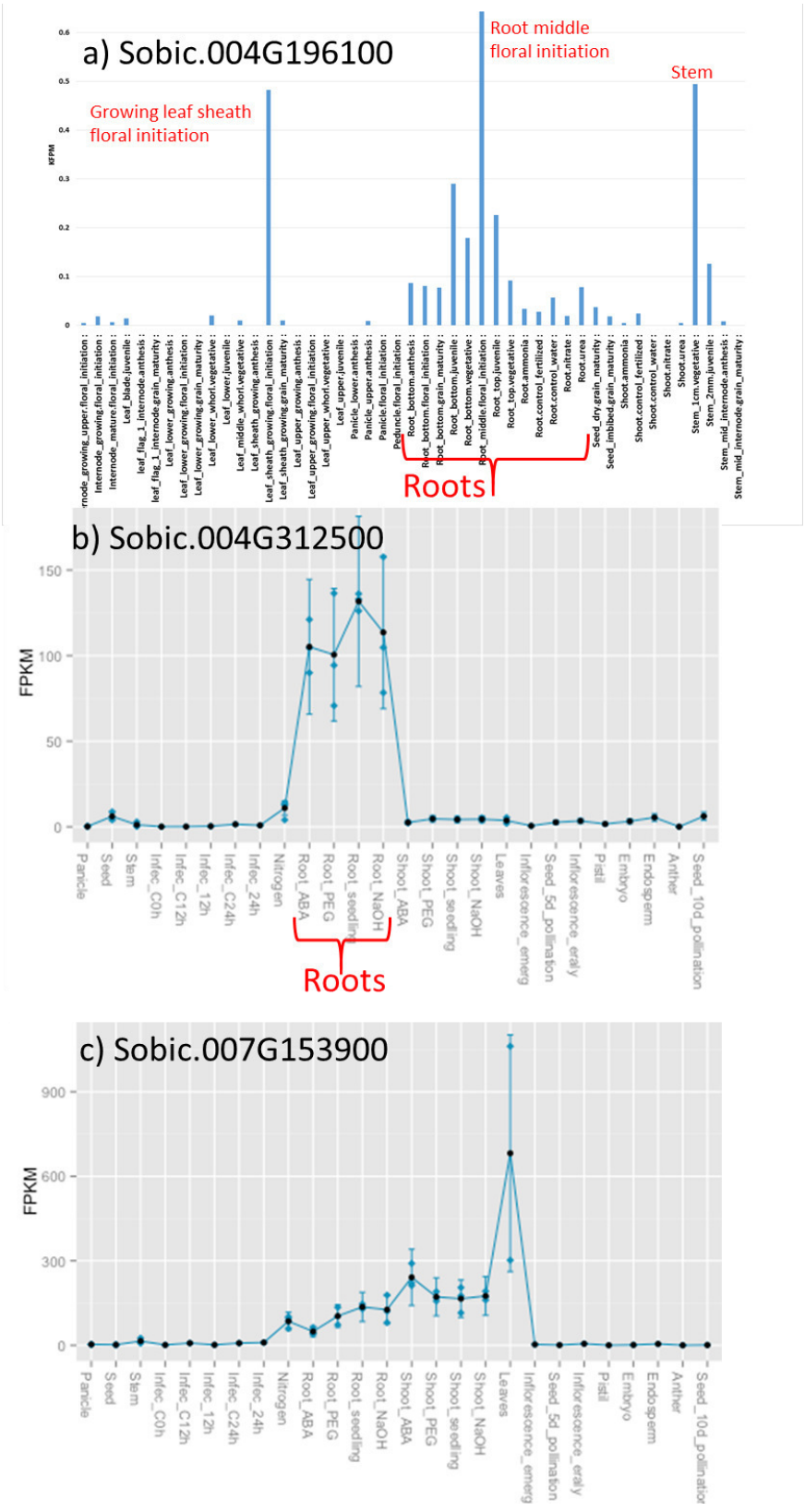

Figure 2.2 Expression patterns of three nitrate reductase genes in sorghum. a) Sobic.004G196100, b) Sobic.004G312500 and c) Sobic.007G153900.

Data from http://sorghum.riken.jp/morokoshi/Home.html (Makita et al. 2015).

FPKM = fragments per kilobase of exon per million fragments mapped.

2.3.2 Sobic.009G225700 – NADH-GOGAT

(gDNA = 11,161 bp with 20 introns; CDS = 6,516 bp; protein = 2,172 amino acids)

There are three genes in sorghum homologous to GOGAT (Figure 2.3; GOGAT = ammonium assimilation enzyme). The sorghum GOGAT homologues are Sobic.009G225700, Sobic.003G258800 and Sobic.002G402700, with only the first being under selection, balancing selection in the Improved and Landrace groups (Table 2.3). The GOGAT genes are separated based on the specificity for electron donors, ferredoxin and NADH (Figure 2.3). The GOGAT gene, Sobic.003G258800, has been shown to be a major part of the ortho-meta QTL for NUE in cereals (Quraishi et al., 2011). The sorghum GOGAT gene under selection (Sobic.009G225700) belongs to a group including the rice NADH-GOGAT2 gene (Figure 2.3) which has been shown to be involved in spikelet number (Tamura et al., 2011). Rice knock-out mutants for the NADH-GOGAT2 showed a significant decrease of 26-39% in spikelet number per panicle associated with a reduction in yield and plant biomass, as well as total N in senescing leaves (Tamura et al., 2011). The rice NADH-GOGAT2 gene is expressed mainly in mature leaves and leaf sheaths, GUS staining in the phloem parenchyma cells and phloem companion cells of vascular bundles (Tamura et al., 2011 and references within). The other rice NADH-GOGAT gene, NADH-GOGAT1 (Os01g48960; Figure 2.3), is mainly expressed in growing tissue such as root tips, young spikelets and developing lead blades, and is important for N remobilization.

The gene Sobic.009G225700 (under selection) is mainly expressed in reproductive tissues (Figure 2.4). The expression of Sobic.003G258800 is mainly in both roots and shoots, whereas the expression of Sobic.002G402700 is highest mainly only in shoots. In general, the relative expression levels of the three genes is; Sobic.003G258800≈ Sobic.002G402700>> Sobic.009G225700.

The enzymatic assay for GOGAT activity requires the addition of glutamine, 2 oxo-glutarate and either NADH or methyl viologen (ferredoxin substitute) to a protein extract and measuring the glutamate formed and glutamine disappeared using HPLC (Akira Suzuki pers. comm.; Suzuki et al. 1982; Martin et al. 1982; Lea et al., 1990). A major problem in green plant tissue is that NADH-GOGAT activity is relatively low compared to the ferredoxin-dependent activity (Suzuki et al., 1987).

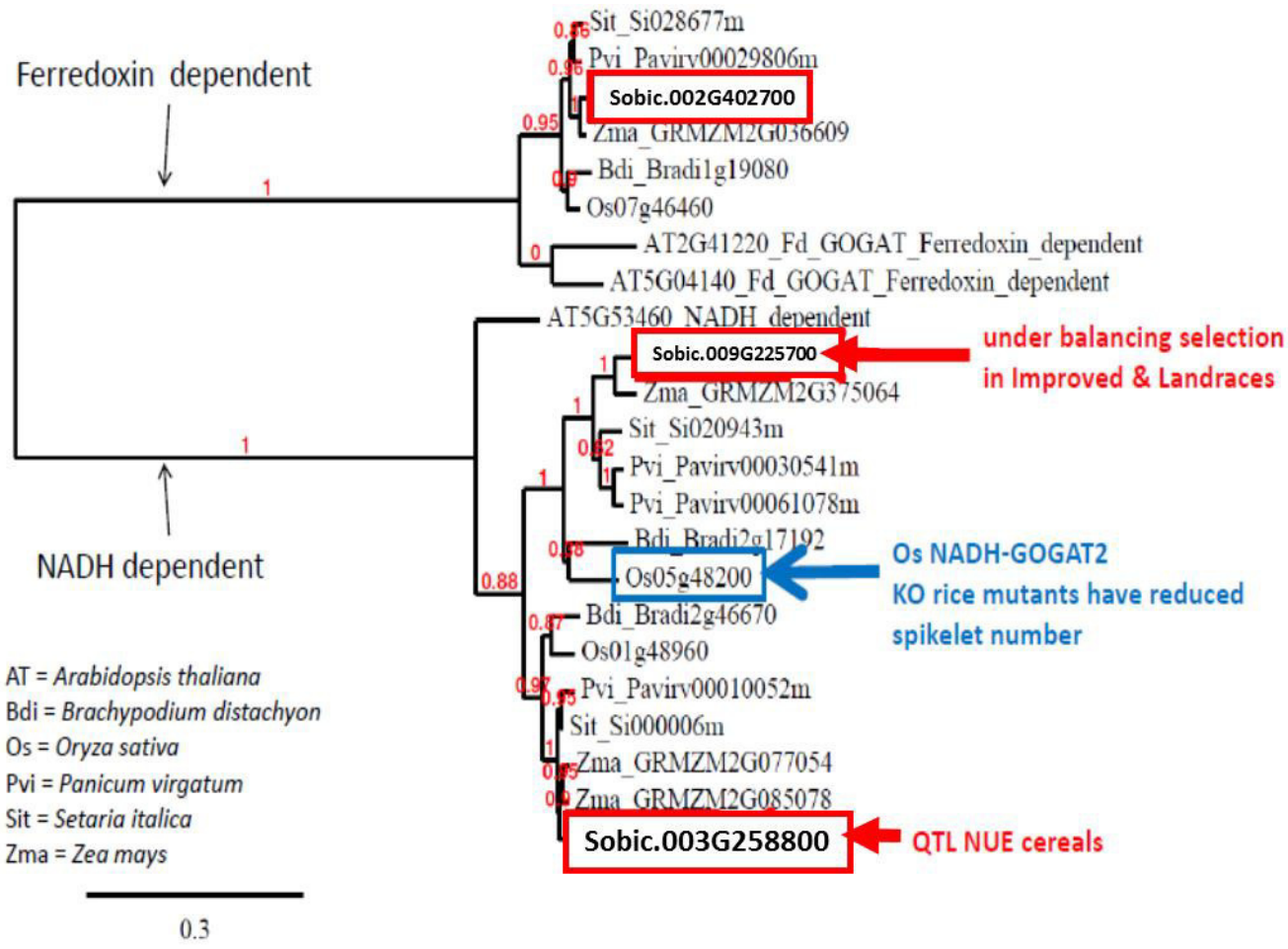

Figure 2.3 Phylogenetic relationship between the protein sequences of GOGAT in *Arabidopsis* and their homologues in the monocots with completed whole genome sequences. (Protein sequences extracted from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) and phylogenetic analyses conducted at http://www.phylogeny.fr (Table 2.1).

Red box = *Sorghum bicolor* GOGAT genes.

Scale bar = 3 amino acid substitutions per 10 amino acids.

Red numbers represent a measure of support for the node; 0-1 where 1 represents maximal support.

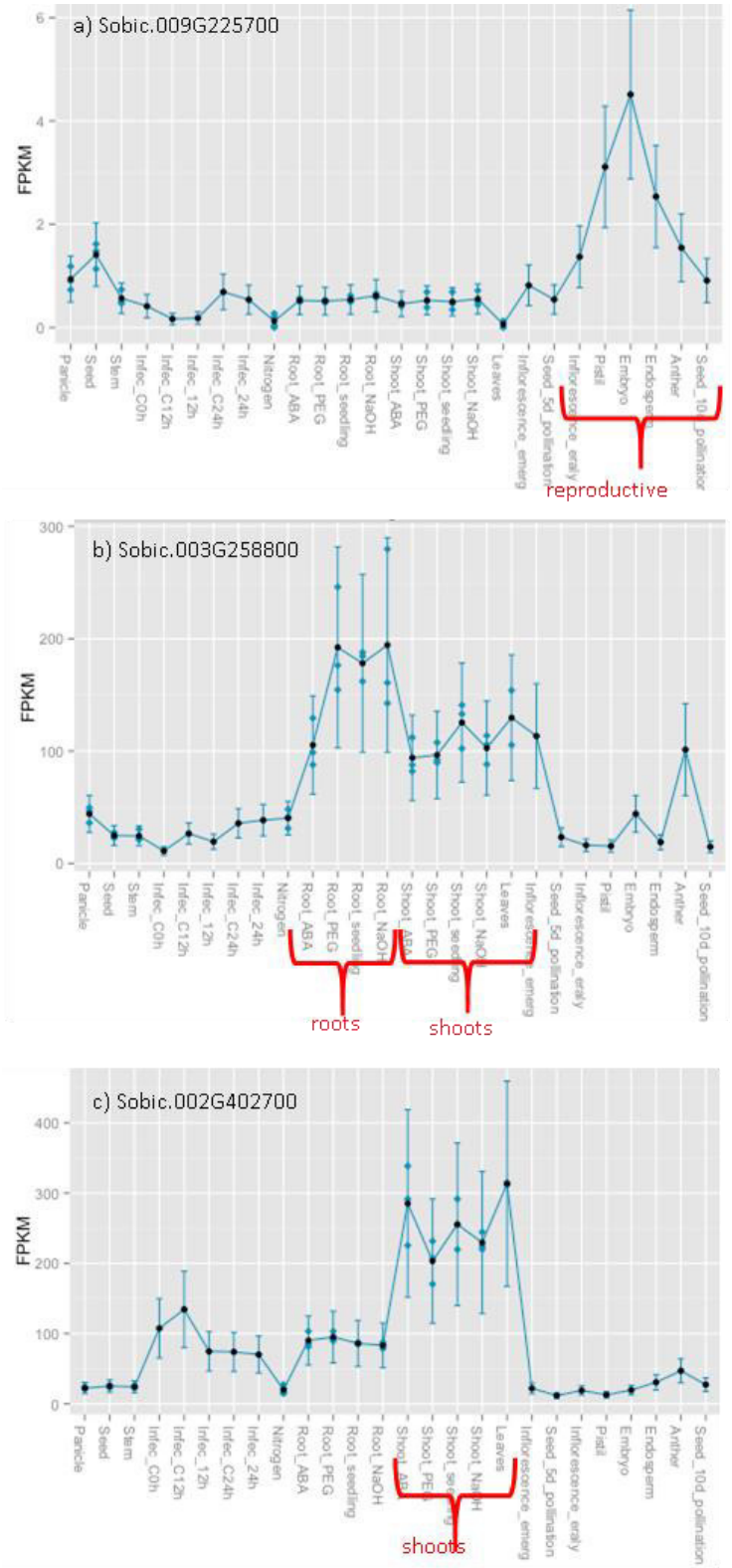

Figure 2.4 Expression patterns of three GOGAT genes in sorghum. a) Sobic.009G225700, b) Sobic.003G258800 and c) Sobic.002G402700.

Data from http://sorghum.riken.jp/morokoshi/Home.html (Makita et al. 2015).

FPKM = fragments per kilobase of exon per million fragments mapped.

2.3.3 Selection of genotypes with contrasting alleles for the two genes

To select genotypes with contrasting alleles for the NR (Sobic.004G196100) and NADH-GOGAT (Sobic.009G225700) genes, phylogenetic analyses were conducted on the predicted protein sequences of these genes for the 44 resequenced sorghum genomes (Figure 2.5). The genotypes were selected based on four criteria:-

1) genotypes are from the Improved category where these two genes have been identified to be under balancing selection (see previous sections),

2) genotypes are parents of NAM populations created by Jordan et al. (2011),

3) molecular marker data is available for the progeny of the NAM populations with the selected genotypes as parents, to enable selection of different alleles of the two genes, and

4) sufficient seed of the parents and progeny is immediately available to conduct a glasshouse experiment without the need for seed multiplication.

From these criteria, three genotypes were selected as parents of NAM populations, R931945-2-2 (recurrent parent), and Macia and ICSV745 (both non-recurrent parents). R931945-2-2 is an Australian elite, midge resistant, highly stay green breeding line (Jordan et al., 2011, 2012). The other two parents (non-recurrent), Macia is a Mozambique cultivar selected for yield and drought tolerance, and ISCV745 was originally selected as a highly midge resistant line at ICRISAT in 1985 (Jordan et al., 2011). Consequently, two NAM populations were further examined based on the crosses, R931945-2-2 x Macia, and R931945-2-2 x ISCV745. Progeny from these two crosses were then analyzed by molecular markers (GBS/DArT/SNP) to identify lines that had the various allelic combinations of the two genes sourced from the two non-recurrent parents (Table 2.4).

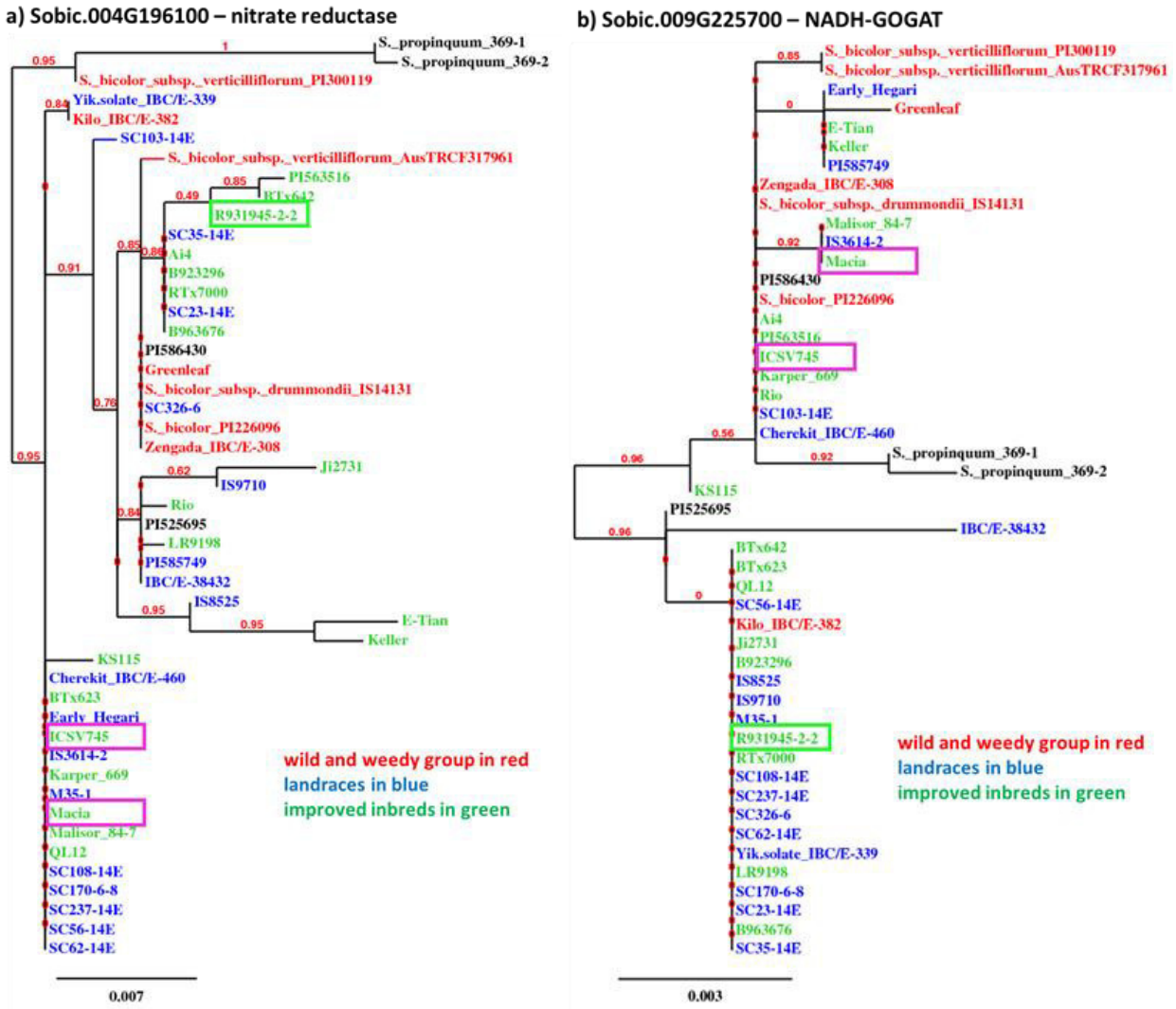

Figure 2.5 Phylogenetic relationship between the predicted protein sequences of the nitrate reductase (Sobic.004G196100) and NADH-GOGAT (Sobic.009G225700) genes that are under balancing selection in the genomes of 44 sorghum genotypes. The predicted protein sequences were obtained by translating the DNA sequences obtained from the 44 genomes using the AUGUSTUS program (Table 2.1) trained to the translation of the two genes.

Scale bar = number of substitutions per site.

Pink box = non-recurrent parents Green box = recurrent parent

Table 2.4 Details of the nine sorghum genotypes.

NAM progeny identification	Recurrent parent	Non-recurrent parent	Source of allele		Percentage of recurrent parent genome (%)
			Nitrate reductase (Sobic.004G196100)	NADH-GOGAT (Sobic.009G225700)	
R03128 – 71	R931945-2-2	ICSV745	ICSV745	R931945-2-2	85
R03128 – 32	R931945-2-2	ICSV745	R931945-2-2	ICSV745	85
R03128 – 66	R931945-2-2	ICSV745	ICSV745	ICSV745	82
R04042 – 25	R931945-2-2	Macia	Macia	R931945-2-2	83
R04042 – 56	R931945-2-2	Macia	R931945-2-2	Macia	79
R04042 – 105	R931945-2-2	Macia	Macia	Macia	80

2.4 Conclusion

In conclusion, two genes, a nitrate reductase and a NADH-GOGAT gene, were identified as both being associated with NUE and also being under selection in the population of 44 sorghum genotypes sequenced. The protein products of these genes are involved in primary nitrate assimilation and glutamate synthesis.

Phylogenetic analyses of both the DNA and amino acid sequences of the two genes in the 44 sorghum genomes sequences, allowed the selection of genotypes with contrasting sequences for these two genes, and also the three genotypes are parents of a recently constructed NAM populations. The use of molecular markers enabled the selection of NAM progeny with varying combinations of alleles of the two genes sourced from the three parents. The following chapter examines the vegetative responses to varying nitrate treatments of these three parents and their progeny from a NAM population.

CHAPTER 3
(manuscript submitted to Molecular Breeding)

ANALYSIS OF THE VEGETATIVE NITROGEN RESPONSE OF SORGHUM LINES CONTAINING DIFFERENT ALLELES FOR NITRATE REDUCTASE AND GLUTAMATE SYNTHASE.

Eugene Diatloff[1*], Emma S. Mace[2], David R. Jordan[3], Sophie Filleur[4,5], Shuaishuai Tai[6], Susanne Schmidt[1] and Ian D. Godwin[1]

* Corresponding author – eugene.diatloff@gmail.com

[1] The University of Queensland, School of Agriculture and Food Sciences, Brisbane, QLD 4072, Australia susanne.schmidt@uq.edu.au i.godwin@uq.edu.au eugene.diatloff@uq.net.au

[2] Department of Agriculture, Fisheries and Forestry, Hermitage Research Facility, Warwick, QLD 4370, Australia emma.mace@daff.qld.gov.au

[3] Queensland Alliance for Agriculture and Food Innovation, Hermitage Research Facility, University of Queensland, Warwick, Qld 4370, Australia david.jordan@uq.edu.au

[4] CNRS I2BC, Department of Cell Biology, Integrated Ion transport group, Bat 21, 1 avenue de la Terrasse, 91190 Gif sur Yvette, France. Sophie.FILLEUR@i2bc.paris-saclay.fr

[5] Université Paris 7 Denis Diderot, U.F.R. Sciences du Vivant, 35 rue Hélène Brion, 75205 Paris Cedex 13, France.

[6] BGI-Shenzhen, Shenzhen 518083, China. Shuaishuai.tai@genomics.cn

Abstract

Improving the nitrogen (N) responsiveness of crops is crucial for food security and environmental sustainability, and breeding N use efficient (NUE) crops has to exploit genetic variation for this complex trait. We used reverse genetics to examine allelic variation in two N metabolism genes. *In silico* analysis of the genomes of 44 genetically diverse sorghum genotypes identified a nitrate reductase and a glutamate synthase gene (NADH-GOGAT) that were under balancing selection in improved sorghum cultivars. We hypothesised that these genes are a potential source of differences in NUE, and selected parents and progeny of nested association mapping populations with different allelic combinations for these genes. Allelic variation was sourced from African (Macia) and Indian (ICSV754) genotypes that had been incorporated into the Australian elite parent R931945-2-2. Nine genotypes were grown for 30 days in a glasshouse and supplied with continuous limiting or replete N, or replete N for 27 days followed by three days N starvation. Biomass, N and nitrate contents were quantified together with gene expressions in leaves, stems and roots. Limiting N supply universally resulted in less shoot and root growth, increased root weight ratio, reduced tissue nitrate and N concentrations, and reduced NADH-GOGAT expression. None of the tested genotypes exceeded growth or NUE of elite parent R931945-2-2 indicating that the allelic combinations did not confer an advantage during early vegetative growth. Thus, the next steps for ascertaining potential effects on NUE include growing plants to maturity. We conclude that reverse genetics that take advantage of rapidly expanding genomic databases contributes towards a systematic approach for developing N efficient crops.

Keywords: allelic variation, sorghum, Nested Association mapping population, nitrogen use efficiency (NUE), vegetative, nitrate reductase, glutamate synthase

Abbreviations: DW dry weight, SA surface area, SLN specific leaf nitrogen, FW fresh weight, N nitrogen, NAM Nested Association Mapping

Introduction

Nitrogen (N) is essential for plant biomass production, grain yield and quality. Nitrogen fertilizer together with plant genetic improvement have majorly contributed to the "Green Revolution" that enabled substantial increases in crop yield (Erisman et al., 2008). Much of the past research on N fertilizer use has addressed questions about the "rates and dates" of fertilizer application, and the success of this approach cannot be denied. However, more recently plant physiologists and breeders have been challenged by concerns that crops systems should have greater N use efficiency (NUE) to reduce costs for growers and environmental pollution (Hirel et al., 2007). A two-tiered approach is considered most effective, composed of improved agronomies and crop cultivars with efficient N uptake and internal N use. NUE as a trait has resisted rapid progress, although breeding programs of major crops are targeting NUE (reviewed by Robinson et al. 2015). The experienced difficulties include uncertainty about effective screening for NUE, considerable genotype x environment x management (GxExM) interactions, and incomplete understanding of the molecular basis underlying NUE.

Here we aim to advance knowledge of the molecular basis of NUE that capitalizes on published genomes of model plant and crop species. This approach interrogates genomic data for allelic variation in genes involved in N acquisition, assimilation or metabolism, and assesses their role in NUE. We focused on sorghum (*Sorghum bicolor* (L.) Moench), a globally important crop which had its first draft genome published by Paterson *et al.* (2009). More recently, the high coverage re-sequenced genomes of 44 sorghum lines have become available, representing the primary gene pool and spanning broad geographic origins, end-uses and taxonomic groupings (Mace *et al.*, 2013). These genomes provide an unmatched resource covering much of the genetic diversity in sorghum and enabling the detailed investigation of genes involved in NUE. This is particularly important since, although there have been some genetic improvements in sorghum, comparatively few resources have been invested into sorghum breeding compared to other cereal crops (Dillon *et al.*, 2007). Sorghum is an important crop in semi-arid regions including many developing countries, and productivity is limited by soil fertility, especially N. Consequently, identification of new sources of genetic variability is essential to assemble new cultivars that are better adapted to resource limitations including N.

Several studies have examined the N responses of a wide range of sorghum genotypes (eg. Maranville *et al.*, 2002 a,b; Youngquist *et al.*, 1992). An alternative and more systematic approach, which has not yet been reported in sorghum is to phenotype genotypes that contain allelic variation

specifically for genes involved in N uptake, assimilation and metabolism in sorghum. Selection of genotypes with allelic variation is now enabled by genetic resources such as the sorghum Nested Association Mapping (NAM) populations created by Jordan et al. (2011), and the availability of genomic sequencing data (Mace et al. 2013). We used these resources to identify two genes, a nitrate reductase and a glutamate synthase gene, which were under balancing selection and are involved in N assimilation and metabolism. We then selected NAM population parents and progeny that contained varying combinations of contrasting alleles for these two genes, and phenotyped these genotypes for their vegetative responses to varying levels of N.

Materials and methods

In silico analyses of published sorghum genomes

The aim of the *in silico* analyses was to identify and select sorghum genotypes containing combinations of contrasting alleles for genes involved in nitrogen assimilation and metabolism that are under selection in the genomes of 44 sorghum genotypes. Population genomic analyses were conducted on the re-sequenced 44 sorghum genomes described by Mace et al. (2013). Analysis was performed by evaluating nucleotide diversity ($\theta\pi$; Nei, 1987), Watterson's estimator ($\theta\omega$; Watterson, 1975), and neutrality test *Tajima's D* (TajD; Tajima, 1989) using BioPerl modules with an in-house script (Mace et al., 2013). Genes under balancing selection were identified and were further analysed for their functional annotation (https://www.blast2go.com/) to identify those potentially involved in N assimilation and metabolism. Genes satisfying these criteria were subjected to phylogenetic analyses (www.phylogeny.fr) of their DNA and predicted amino acid sequences (http://bioinf.uni-greifswald.de/webaugustus/) from each of the 44 sorghum genomes to select those genes with similar phylogenetic relationships in both DNA and amino acid sequences. From all these analyses, two genes were identified (Table 1), and three sorghum genotypes were identified that contained contrasting sequences for the genes of interest (Table 2) and that were also parents of NAM populations recently described by Jordan et al. (2011). Progeny from two such NAM populations were characterized by molecular markers (GBS/DArT/SNP) to select genotypes that had various allelic combinations of the genes of interest. The sorghum genotypes used in this study are listed in Table 2, and their selection and characteristics are described in more detail in the Results section "*In silico* selection of sorghum genotypes".

Table 1. The summary population statistics of two sorghum genes.

***** = Tajima's D values >2 indicating balancing selection.

Gene	Annotation	Selection	Selection criteria								
			Tajima's D			$\theta\pi$ ($x10^{-3}$)			$\theta\omega$ ($x10^{-3}$)		
			Landraces	W & W	Improved	Landraces	W & W	Improved	Landraces	W & W	Improved
Sobic.004G196100	Nitrate reductase	Balancing in Improved	1.3	0.8	2.1*	2.6	3.7	3.3	1.7	2.5	1.8
Sobic.009G225700	NADH-GOGAT	Balancing in Improved & Landraces	3.0*	-1.2	2.2*	3.7	2.5	3.5	2.0	2.5	1.9

W & W = Wild and weedy

See Mace et al. (2013) for genotype classification

Table 2 Details of the nine sorghum genotypes.

NAM progeny identification	Recurrent parent	Non-recurrent parent	Source of allele		Percentage of recurrent parent genome (%)
			Nitrate reductase (Sobic.004G196100)	NADH-GOGAT (Sobic.009G225700)	
R03128 – 71	R931945-2-2	ICSV745	ICSV745	R931945-2-2	85
R03128 – 32	R931945-2-2	ICSV745	R931945-2-2	ICSV745	85
R03128 – 66	R931945-2-2	ICSV745	ICSV745	ICSV745	82
R04042 – 25	R931945-2-2	Macia	Macia	R931945-2-2	83
R04042 – 56	R931945-2-2	Macia	R931945-2-2	Macia	79
R04042 – 105	R931945-2-2	Macia	Macia	Macia	80

Growth conditions

To evaluate the effects of allelic variation in early vegetative growth, sorghum plants were grown for 30 days in a glasshouse at constant 25°C during August 2014, Brisbane, Australia (day length was between 10h 49 min and 11h 30 min). The average midday light intensity at the leaf canopy level was 660±60 µmol/m^2/s. Seeds were sown in 137 mm top diameter pots with saucers (140 mm tall; http://www.anovapot.com/) containing washed river sand (pH 6.5). The experimental design was nine sorghum genotypes x three N rates x three replicates (81 pots). Each pot contained three plants. Plants were watered with a nutrient solution every 2-3 days. The three N treatments were: continuous replete (high) N (10 mM nitrate), continuous limiting (low) N (1 mM nitrate), and high N (10 mM nitrate) for 27 days, followed by no N for the last three days prior to harvest. Each pot received 50 mL of nutrient solution: 2 mM $MgSO_4$; 2 mM $CaSO_4$; 0.457 mM KH_2PO_4; 42.5 µM K_2HPO_4; 100 µM FeEDTA; 10 µM $MnSO_4$; 10 µM H_3BO_3; 1 µM $CuSO_4$; 2.5 µM $ZnSO_4$; and 0.35 µM Na_2MoO_4 (Robinson et al. 2011). One of three nitrate treatments: 1mM KNO_3 (+ 4.5mM K_2SO_4); or 10mM KNO_3 were added to each nutrient solution. At day 27, one of the 10 mM KNO_3 treatments was replaced with 5 mM K_2SO_4 in the N starvation treatment.

Tissue sampling, growth measurements and nitrogen analysis

At harvest (0900 to 1130h), the two most uniform plants in each pot were selected and separately dedicated to either gene expression and nitrate analysis, or plant growth and total N quantification. Each plant was separated into washed roots, leaves (leaf blades) and stems (leaf sheaths). The separated material from one plant was frozen in liquid N_2 and stored at -80°C for

total RNA extraction, cDNA synthesis and qPCR analysis for gene expression, and nitrate determination. The material from the other plant was subjected to root fresh weight measurement, total leaf area measurement (LI-3100C Area Meter, Licor, USA), followed by drying at 60°C for four days prior to dry weight measurement and total N analysis. Total N was determined in ground samples by dry combustion and infrared detection in a LECO analyser (CNS-2000, LECO Corporation, MI, USA). Nitrate contents were analysed on frozen samples after water extraction (Miranda *et al.*, 2001).

Total RNA extraction and amplification of gene transcript

Frozen plant material was ground in a mortar and pestle containing liquid N_2 and total RNA extracted using the ISOLATE II RNA Plant Kit according to the manufacturer's instructions (www.bioline.com/). Total RNA was quantified and assessed for purity on a Nanodrop (www.nanodrop.com/).

The aim of the qPCR analyses was to examine the expression of the nitrate reductase and NADH-GOGAT genes under selection (Table 2) and also of the other family members of these two genes, in the three different tissues types from the nine genotypes grown under three N levels. A search of the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html#) indicated that the sorghum genome contains three nitrate reductase homologues and three GOGAT homologues. We were unable to successfully design gene specific qPCR primers for each of the three NR genes to discriminate between them due to their close 72-81% DNA identity. Gene specific primers for the three GOGAT genes were successfully designed and tested (Table 1S in Appendix), and consequently only the expression of these three GOGAT genes was investigated in the phenotyping experiment.

qPCR was conducted on a Roche Lightcycler 480 using the SensiFAST SYBR No-ROX One-Step-Kit (www.bioline.com/) in 10 µL reactions in 96 well plates according to the manufacturers' instructions. The PCR program was initiated with a preincubation step of 45°C for 200 sec followed by 95°C for 120 sec, then the 3 step amplification by 55 thermal cycles of 95°C for 5 sec, 60°C for 10 sec, and 72°C for 5 sec. Melting curve analysis was performed immediately after the real-time PCR. The temperature range used for the melting curve generation was from 65°C to 97°C. Each 96-well plate was a balanced design containing genotypes and nitrate treatment biological replicates for a single tissue type and primers for the three GOGAT genes and two reference genes. Each 96-well plate contained template-free (water) and reverse transcriptase-free

controls. LinRegPCR software was used to determine PCR efficiency and calculate starting concentration of mRNA, N_0 (Ruijter *et al.*, 2009). Relative gene expression was determined by N_0 (gene of interest) / N_0 (reference genes). Two reference genes were used (Sobic.009G174900 - similar to Ubiquitin carrier protein; Sobic.002G339600 - similar to Elongation factor 1-alpha). The denominator of the relative gene expression ratio was the geometric mean of the N_0 for both reference genes as per Vandesompele et al. (2002). The primers used are listed in Table 1S (Appendix).

Calculations of dry matter and nitrogen relations

The root weight ratio was calculated as the ratio of root DW to total plant DW. Specific leaf nitrogen (SLN) was calculated as the amount (g) of N per m² of leaf area (LA). Net N uptake was calculated according to Williams (1948):-

$$V = \frac{(lnWR2 - lnWR1)(M2 - M1)}{(t2 - t1)(WR2 - WR1)}$$

where V is the mean rate of N absorption per unit fresh weight (FW) of root (µg of N/mg of root FW/day), WR1 and WR2 are the initial seed dry weight (mg) and fresh root weight (mg) at harvest, M1 and M2 the total N contents (µg of N) in the seed (t1=day 0) and at harvest (t2=day 30). The assumptions underlying this and similar formulae have been described fully elsewhere (Williams 1948; Silberbush and Gbur 1994). The total plant and shoot N usage indices were calculated according to the modified utilisation index (UI) proposed by (Siddiqi and Glass, 1981):-

$UI = DW^2.N$

where DW and N are the dry weights (g) and N contents (g of N) of either the total plant or shoots (leaf (leaf blade) plus stem (leaf sheath)).

Statistical analysis

The experimental design was a split-plot with genotypes as the main plot and N rates as subplots. Results were analysed by GenStat (17th edition, https://www.vsni.co.uk/software/genstat/) and means compared by the Tukey's HSD test ($p \leq 0.05$).

Results

In silico selection of sorghum genotypes

In silico analyses of the 44 sorghum genomes identified two genes (nitrate reductase (NR - Sobic.004G196100) and a glutamate synthase (NADH-GOGAT - Sobic.009G225700)) that were under balancing selection in improved inbred and landrace sorghum lines. The genes contained functional annotation indicating their likely involvement in N assimilation and metabolism, and also showed similar phylogenetic relationships at both the DNA and amino acid levels among the 44 sorghum genomes (Table 1). The latter observation indicates that differences at the DNA level were translated into differences in the amino acid sequences, and potentially influencing the function of these proteins and subsequent plant phenotype. Balancing selection maintains genetic diversity within populations (Delph and Kelly, 2014; Wright and Gaut, 2005). These polymorphisms are more likely to be segregating at intermediate frequencies, where they contribute most to population variance affecting fitness and consequently there are good reasons to be interested in identifying balanced polymorphisms in a species (Tian *et al.*, 2002).

Genotypes were selected containing varying combinations of different alleles for the two genes, based on four criteria:-

1) parental genotypes were from the Improved category where these two genes were identified to be under balancing selection (Table 1),

2) parental genotypes were parents of NAM populations (Jordan et al. 2011),

3) molecular marker data were available for the progenies of the NAM populations with the selected genotypes as parents, to enable selection of different alleles of the two genes, and

4) sufficient seed of the parents and progeny was immediately available to conduct a glasshouse experiment without the need for seed multiplication.

From these criteria, three genotypes were selected as parents of NAM populations, R931945-2-2 (recurrent parent), and Macia and ICSV745 (both non-recurrent parents) (Table 2). R931945-2-2 is an Australian elite, midge resistant, highly stay green breeding line (Jordan et al., 2011, 2012). The other two parents (non-recurrent), Macia is a Mozambique cultivar selected for

yield and drought tolerance, and ISCV745 was originally selected as a highly midge resistant line at ICRISAT in 1985 (Jordan et al., 2011). Consequently, two NAM populations were further examined based on the crosses, R931945-2-2 x Macia, and R931945-2-2 x ISCV745. Progenies from these two crosses were then analyzed by molecular markers (GBS/DArT/SNP) to identify lines that had the various allelic combinations of the two genes sourced from the two non-recurrent parents (Table 2). Consequently, nine sorghum genotypes were used for phenotyping under a range of N conditions. Where significant ($P<0.05$) genotype effects are detected, comparisons of means are made mainly with those of the recurrent genotype R931945-2-2, since it is the recurrent parent (79-85% genetic background of NAM progeny, Table 2) and is considered the elite breeding line.

Plant dry weight parameters

Nitrogen supply had a highly significant ($P<0.001$) effect on all sorghum growth parameters measured (Table 3). Decreasing N concentrations from replete (high 10 mM nitrate) to limiting (low 1 mM nitrate), significantly affected shoot parameters most strongly with 52, 46 and 59% reductions in leaf DW, stem DW and leaf area, respectively. Root DW was less affected and reduced by 26%, and total DW in the high-N followed by three days N starvation was significantly lower than in the 10 mM treatment, except root DW, which was similar in both treatments. No significant differences in root weight ratio were observed between high-N+starved and high N treatments, but leaf SA was significantly lower in the high N and N starvation than in the 10 mM treatment. There were no significant genotype effects on leaf DW, stem DW and LA, but highly significant ($P<0.001$) genotype effects were detected in root DW, and significant ($0.05 \geq P > 0.01$) genotype effects were detected in total DW and root weight ratio. The recurrent parent R931945-2-2 produced significantly higher root DW than the non-recurrent parents, Macia and ICSV745 (Table 4). None of the six NAM population progenies produced higher root DW than R931945-2-2. Genotype ranking observed for root DW was the same for root fresh weight (data not shown). The genotype R931945-2-2 (recurrent parent) produced significantly higher root weight ratio than the two non-recurrent parents, Macia and ICSV745 (Table 4). None of the six NAM population progeny produced higher root weight ratios than R931945-2-2. There were no significant differences between the genotypes in their total DW, even though a significant F-test was obtained (Tables 3 & 4). This most likely due to the different methodologies used to calculate the p-value from an ANOVA and a post-hoc analysis. The Tukey's HSD method is more conservative that the F-test, which may explain the difference in the results. There were no significant ($P>0.05$) genotype x N interactions in any of the measured DW parameters (Table 3).

Table 3 Phenotypic values of nine sorghum genotypes grown at three nitrate levels.

Level of significance is indicated by; . for not significant where $P>0.05$, * for $0.05 \geq P > 0.01$, ** for $P<0.005$ and *** for $P<0.001$.

Values presented for the means for the nitrate effect are means across all genotypes. Means within a row with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test.

				Means for nitrate effect		
Parameter	Genotype (G)	Nitrate (N)	G x N	1 mM	10 mM starved	10 mM
Leaf DW (mg/plant)	.	***	.	110a	187b	228c
Stem DW (mg/plant)	.	***	.	59a	95b	110c
Root DW (mg/plant)	***	***	.	78a	95b	105b
Total DW (mg/plant)	*	***	.	248a	377b	443c
Root weight ratio	*	***	.	0.31a	0.25b	0.24b
Leaf area (cm^2/plant)	.	***	.	55a	114b	134c
Total N in Leaf (%)	*	***	.	2.3a	4.1b	5.0c
Total N in Stem (%)	*	***	.	1.2a	3.7b	5.1c
Total N in Root (%)	.	***	.	1.3a	2.4b	2.7b
SLN (g of N/m^2 leaf)	.	***	.	0.45a	0.67b	0.85c
Nitrate-N in Leaf (mg N/kg DW)	*	***	.	6a	313b	788c
Nitrate-N in Stem (mg N/kg DW)	*	***	.	7a	1748b	2870c
Nitrate-N in Root (mg N/kg DW)	**	***	.	12a	410b	1047c
Net N uptake (µg of N/mg of Root FW/day)	*	***	.	1.31a	2.71b	3.43c
Total Plant Usage Index (g DW^2/g of N)	.	***	.	16.7a	10.9b	9.9b
Shoot Usage Index (g DW^2/g of N)	.	***	.	10.3a	7.4b	6.7b

Table 4 Genotype effects on plant physical parameters and ratios.

Genotypic means (ie. means across all nitrate levels) are presented for those parameters where a significant genotype effect was detected in Table 3.

Means within a column with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test.

RP = recurrent parent, NRP = non-recurrent parent.

	Genotype	Root DW (mg/plant)		Genotype	Root weight ratio		Genotype	Total plant DW (mg/plant)
RP	R931945-2-2	124a	RP	R931945-2-2	0.30a	RP	R931945-2-2	421a
	R04042-25	107ab		R04042-25	0.28ab		R04042-105	399a
	R04042-105	102ab		R03128-71	0.27ab		R04042-25	386a
	R03128-32	98abc		R03128-32	0.27ab		R03128-32	370a
	R03128-71	88bc		R03128-66	0.27ab	NRP	Macia	351a
NRP	Macia	86bc		R04042-105	0.26ab		R03128-66	331a
	R03128-66	85bc	NRP	Macia	0.25b		R03128-71	329a
	R04042-56	73c		R04042-56	0.25b	NRP	ICSV745	308a
NRP	ICSV745	73c	NRP	ICSV745	0.24b		R04042-56	307a

Tissue nitrogen and nitrate concentrations

Nitrogen supply had a highly significant (P<0.001) effect on all N% and nitrate-N measurements (Table 3) with lower values in the low compared to high N treatments. All parameters were significantly lower in the high N+starvation treatment than in the high N treatment, except for N% in roots which were similar in both high-N treatments. Significant ($0.05 \geq P > 0.01$) genotype effects were detected in N% and nitrate concentrations of leaf and stem, and highly significant (P<0.0.01) genotype effects occurred in root nitrate concentrations (Table 3). No genotype had significantly different higher or lower values than R931945-2-2 for N% in stem or nitrate concentrations in leaf or stem (Table 5). Only genotype R03128-71 contained significantly (P<0.05) higher leaf N% and root nitrate concentrations than R931945-2-2. There were no significant (P>0.05) genotype x N interactions in the measured N% or nitrate concentrations for the three plant tissues (Table 3).

N uptake and usage index

The concentration of nitrate in the root zone had a highly significant (P<0.001) effect on the net N uptake (Table 3). Net N uptake in the low N treatment was significantly lower than that in the high N+starvation treatment, which in turn, recorded significantly lower net N uptake than in the high N treatment (Table 3). There was a significant ($0.05 \geq P > 0.01$) genotype effect on net N uptake (Table 3). The net N uptake of genotypes R03128-66 (NAM progeny) and ICSV745 (non-recurrent parent) was significantly higher than that of R931945-2-2 (Table 6), and no genotypes showed a net N uptake rate significantly lower than R931945-2-2 (Table 6).

The concentration of nitrate in the root zone had a highly significant (P<0.001) effect on the two N usage indices (Table 3). Both indices were significantly higher in the low N treatment than in both high N treatments. Starvation of plants of N for three days did not alter usage index compared to the continuous high N treatment and no genotype effects were observed. There were no significant (P>0.05) genotype x N interactions in the net N uptake, nor in either of the N usage indices (Table 3).

Table 5 Genotype effects on the concentrations of total N and nitrate-N in plant tissues.

Genotypic means, ie. means across all nitrate levels, are presented for those parameters where a significant genotype effect was detected in Table 3. Means within a column with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test.
RP = recurrent parent, NRP = non-recurrent parent.

	Genotype	Leaf N (%)		Genotype	Stem N (%)		Genotype	Leaf nitrate-N (mg/kg)		Genotype	Stem nitrate-N (mg/kg)		Genotype	Root nitrate-N (mg/kg)
	R03128-71	4.17a		R03128-71	3.60a		R04042-25	506a	NRP	Macia	1748a		R03128-71	750a
	R04042-56	3.95ab		R04042-25	3.57a		R03128-71	409ab		R04042-25	1746a		R04042-25	547ab
	R04042-25	3.92ab		R04042-56	3.49ab	NRP	Macia	384ab		R04042-105	1738a		R04042-105	510b
NRP	Macia	3.82ab	NRP	Macia	3.45ab	RP	R931945-2-2	370ab	RP	R931945-2-2	1671ab		R04042-56	494b
	R03128-32	3.82ab		R04042-105	3.31ab		R04042-105	364ab		R04042-56	1587ab	RP	R931945-2-2	450b
	R04042-105	3.81ab		R03128-32	3.29ab		R04042-56	363ab		R03128-71	1578ab	NRP	ICSV745	434b
	R03128-66	3.63b	RP	R931945-2-2	3.25ab	NRP	ICSV745	324ab		R03128-66	1345ab	NRP	Macia	415b
RP	R931945-2-2	3.63b	NRP	ICSV745	3.23ab		R03128-32	314b		R03128-32	1340ab		R03128-66	406b
NRP	ICSV745	3.63b		R03128-66	2.98b		R03128-66	290b	NRP	ICSV745	1123b		R03128-32	402b

Table 6 Genotype effects on the net N uptake.

Genotypic means, ie. means across all nitrate levels, are presented for those parameters where a significant genotype effect was detected in Table 3.
Means within a column with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test.
RP = recurrent parent, NRP = non-recurrent parent.

	Genotype	Net N uptake (µg of N/mg root FW/day)
	R03128-66	2.74a
NRP	ICSV745	2.73a
NRP	Macia	2.64ab
	R04042-56	2.58ab
	R04042-25	2.43ab
	R04042-105	2.43ab
	R03128-32	2.40ab
	R03128-71	2.37ab
RP	R931945-2-2	2.04b

GOGAT gene expression

ANOVA was conducted on the four-way interaction (genotype x N x gene x tissue) (Table 7). Similar results were obtained when conducting the ANOVA on the expression of each gene separately (i.e. genotype x N x tissue).

Nitrogen supply had a highly significant (P<0.001) effect on the expression of the three GOGAT genes (Table 7). The mean GOGAT expression across all genotype, gene and tissue treatments was highest in the high N treatment, followed by the high N+starvation treatment, and lowest in the low N treatment (Table 8A). A highly significant (P<0.001) gene effect was detected (Table 7). The mean GOGAT expression across all genotype, N and tissue treatments was highest for Sobic.002G402700, followed by Sobic.003G258800, with very low expression levels detected for Sobic.009G225700, the gene under balancing selection (Table 8B). A highly significant (P<0.001) effect of tissue type was detected (Table 7). The mean GOGAT expression across all genotype, N and gene treatments was highest in the stems, followed by in the roots, and lowest in the leaf tissues (Table 8C).

There were no significant genotype ($P>0.05$) effects on gene expression (Table 7), however, highly significant ($P<001$) genotype x tissue interactions and genotype x gene x N interactions, and significant ($P<0.05$) genotype x gene were detected (Table 7).

The expression of the gene Sobic.009G225700 (under balancing selection) was lowest of all three GOGAT genes in all the three plant tissues and was not significantly affected by N treatment nor by the type of plant tissue (Table 9). The highest expression of all the three genes was for the expression of Sobic.002G402700 in leaf tissues in the high N treatment (Table 9), and this expression was significantly decreased by two days of nitrate starvation and even further reduced by continuous low N treatment (Table 9). The expression of Sobic.003G258800 was induced by the two day nitrate starvation treatment (compared to the high N treatment) in the roots, unaffected by this N treatment in the stems, but decreased by this N treatment in leaf tissues. In general, the expression of the gene Sobic.003G258800 was intermediate between the other two genes.

The expression of Sobic.009G225700 (under balancing selection) did not significantly ($P>0.05$) vary among genotypes (Table 7 and 10). The expression of Sobic.003G258800 was lowest in the three parents and did not significantly differ among these three genotypes (Table 10). The expression of Sobic.003G258800 in the three genotypes R03128-66, R03128-71 and R04042-105 was significantly higher than in two of the three parents, Macia and R931945-2-2. The highest expression of all three genes was for Sobic.002G402700 in the genotype Macia, and this expression was higher than for all other genotypes.

Table 7 ANOVA analysis for the expression values of three GOGAT genes in three different tissues in nine sorghum genotypes grown at three nitrate levels.

Level of significance is indicated by . for not significant where $P>0.05$, * for $P<0.05$, ** for $P<0.005$ and *** for $P<0.001$.

Source	*P* Value	Source (2 way interaction)	*P* value	Source (3 way interaction)	*P* value	Source (4 way interaction)	*P* value
Genotype (G)	.	GxN	.	GxNxGn	.	GxNxGnxT	.
Nitrogen (N)	***	GxGn	*	GxNxT	.		
Gene (Gn)	***	GxT	***	GxGnxT	.		
Tissue (T)	***			NxGnxT	***		
		NxGn	***				
		NxT	***				
		GnxT	***				

Table 8 Effects on the expression of three GOGAT genes in nine sorghum genotypes grown under three nitrate levels.

Means are presented for those parameters where a significant single factor effect was detected in Table 7.

Means with different lowercase letters were significantly different ($P\leq0.05$) according to the Tukey's HSD test.

* = GOGAT under balancing selection.

A

Means for Nitrate effect[1]	
Nitrate concentration	Gene expression
10 mM	1.16a
10 mM starved	1.04b
1 mM	0.69c

[1] = means across genotype, gene and tissue

B

Means for Gene effect[2]	
Gene	Gene expression
Sobic.002G402700	1.76a
Sobic.003G258800	1.08b
Sobic.009G225700*	0.06c

[2] = means across genotype, tissue and N level

C

Means for Tissue effect[3]	
Tissue	Gene expression
Stem	0.078a
Root	0.043b
Leaf	0.031c

[3] = means across genotype, gene and N level

Table 9 Heat map of GOGAT gene expression for nitrogen x gene x tissue interaction (Table 7).
Means are across all genotypes. Means with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test for the 3-way interaction.
* = GOGAT under balancing selection.

	Leaf						Stem						Root					
Gene	1 mM		10 mM starved		10 mM		1 mM		10 mM starved		10 mM		1 mM		10 mM starved		10 mM	
Sobic.009G225700*	0.07	a	0.03	a	0.03	a	0.07	a	0.09	a	0.08	a	0.04	a	0.04	a	0.06	a
Sobic.003G258800	0.21	a	0.56	b	0.87	c	1.34	def	1.52	efghi	1.46	defgh	0.85	c	1.70	ghij	1.23	d
Sobic.002G402700	1.25	de	2.24	k	3.12	l	0.85	c	1.42	defg	1.77	ij	1.56	fghi	1.72	hij	1.86	j

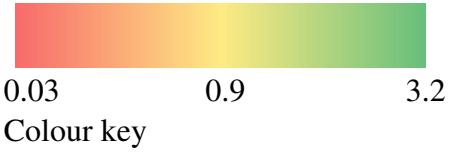

Table 10 Heat map of GOGAT gene expression for genotype x gene interaction (Table 7).
Means are across all tissue types and nitrate treatments. Means with different lowercase letters were significantly different ($P \leq 0.05$) according to the Tukey's HSD test.
RP = recurrent parent, NRP = non-recurrent parent.
* = GOGAT under balancing selection.

	Genotype	Sobic.009G225700*			Genotype	Sobic.003G258800			Genotype	Sobic.002G402700	
NRP	ICSV745	0.07	a		R03128-66	1.29	de	NRP	Macia	2.36	i
	R03128-32	0.07	a		R03128-71	1.26	cde	NRP	ICSV745	1.97	h
	R03128-71	0.07	a		R04042-105	1.15	cde		R03128-71	1.81	gh
	R04042-105	0.06	a		R04042-56	1.11	bcd		R03128-66	1.71	fgh
	R03128-66	0.05	a		R03128-32	1.08	bcd	RP	R931945-2-2	1.68	fgh
	R04042-25	0.05	a		R04042-25	1.05	bcd		R04042-105	1.65	fg
	R04042-56	0.05	a	NRP	ICSV745	1.01	bcd		R03128-32	1.64	fg
RP	R931945-2-2	0.05	a	NRP	Macia	0.97	bc		R04042-25	1.50	ef
NRP	Macia	0.04	a	RP	R931945-2-2	0.84	b		R04042-56	1.48	ef

0.04 1 2.4
Colour key

Discussion

The experiment described here examined the responses of nine sorghum genotypes during early vegetative growth to different N supply to test if certain gene combinations confer higher NUE. The genotypes contained different allelic combinations for two genes involved in N metabolism, a nitrate reductase (NR - Sobic.004G196101) and a glutamate synthase (NADH-GOGAT - Sobic.009G225700) gene, both of which were under balancing selection in improved and landrace sorghum lines (Table 1, Massel *et al.*, 2016). Balancing selection indicates long lived polymorphisms that maintain diversity at allelic sites and can underlie adaptation to environment stresses.

Nitrate reductase is responsible for the first step of nitrate assimilation, the reduction nitrate to nitrite in the cytoplasm. This nitrite is then further reduced in the chloroplast to ammonium. This ammonium is then incorporated into the amino acids, glutamine and glutamate, glutamate synthase being the primary enzyme, in concert with glutamine synthetase (Suzuki and Knaff, 2005). The sources of the natural allelic variation in genes coding for NR and NADH-GOGAT were from three improved sorghum genotypes. Progeny from crosses of these parental lines, containing varying combinations of the NR and NADH-GOGAT genes were selected. The aim of the experiment was to identify if any of the allelic combinations for these two genes from three different parents, altered the nitrate responses. This is the first time that the nitrate responses of the parents and their progenies of sorghum NAM populations have been reported.

Growth - genotypic differences

The growth responses of the nine genotypes to different N supply rate indicated that there was no significant genotype effect on shoot growth (leaf and stem dry weights, leaf area, Table 3). This indicates that the incorporation of allelic variation for NR and GOGAT from the two non-recurrent parents (ICSV745 and Macia) into the genome of the recurrent parent (R931945-2-2) did not confer any significant advantage in shoot growth to the recurrent parent during the early vegetative stage under the conditions of the experiment.

There was however, a highly significant ($P<0.001$) genotype effect on root growth (dry weight, Table 3; and fresh weight, data not shown). None of the genotypes showed significantly higher root DW than R931945-2-2, indicating again that the allelic combinations did not confer any growth advantages. However, the root dry weight of R931945-2-2, was significantly higher

than for either of the two non-recurrent parents (44% and 71% higher than for Macia and ICSV745 respectively). This is in contrast to Singh et al. (2011), who studied the root architecture of 44 inbred and 30 hybrid sorghum genotypes and found that the root dry weight of R931945-2-2 was one of the lowest measured, at a similar early vegetative growth stage of six fully expanded leaves (complete raw data for all genotypes was not presented). It is important to note that the higher root dry weight of R931945-2-2 compared to the other two non-recurrent parental genotypes in the present experiment was obtained in relatively small pots (1.4L pots, containing three plants grown for 30 days). It has been shown that pot volume (7L versus 14, 18, 28 and 56L) can restrict sorghum shoot dry weight, but not root dry weight (Yang *et al.* 2010; 1 plant per pot grown to maturity). Consequently, to confirm the increased root dry weight measured for R931945-2-2 in the present study (and also the non-significant genotype effect on shoot growth), would require experiments using larger soil:root volume ratios. For wheat, it has been shown that genotypes with high early vigour (and root growth) have improved efficiency of N use for biomass production, in addition to improving N uptake during early growth (Pang *et al.*, 2014).

Growth - effect of nitrogen supply

Nitrogen supply had a highly significant ($P<0.001$) effect on all the sorghum growth parameters measured (Table3). Decreasing the root zone nitrate concentration from 10 to 1 mM decreased shoot parameters (leaf DW -52%, stem DW -46%, leaf area -59%) more than root parameters (DW -26%; Table 3). Shoot growth is well known to be more sensitive to reduced concentrations of root zone nitrate than root growth. When starved of N, plants expand their root system at the expense of shoot growth. Since nitrogen is acquired solely by the root system, and an increase in the root weight ratio and root DW (Table 3) reflects resource allocation to roots to enhance root foraging ability (Reynolds and Dantonio, 1996). The short term deprivation of 27 day old sorghum plants of nitrate for three days before harvest, significantly reduced leaf area (-15%), and also the leaf (-18%) stem (-14%) and total (-15%) dry weight, but not that of roots (Table 3). This again emphasises the greater sensitivity of shoot growth, even to short term nitrate deprivation. Reduction in N supply to the roots can reduce leaf expansion within 24 h (Palmer *et al.*, 1996), although how this occurs is not clear (Dodd *et al.*, 2002).

Tissue nitrogen and nitrate concentrations

The N concentrations in the leaves and stems of sorghum plants in the high N treatment were within the adequate range for sorghum of equivalent vegetative growth stage (approx. 5% N;

Table 3; (Grundon et al., 1987; Reuter and Robinson, 1997), but decreased to marginal and deficient concentrations in the low N treatment (<2.5% N; Table 3). This indicates that these two nitrate treatments were suitable for investigating genotypic differences in tissue nitrate and N concentrations under both adequate and marginal/deficient N conditions. The two genes for which allelic variation was investigated in this study are involved in nitrate assimilation and amino acid synthesis, and consequently may result in genotypic differences in the concentrations of nitrate and N in plant tissues. However, the only significant genotype differences were that R03128-71 had significantly higher leaf N% and root nitrate-N than R931945-2-2 (Table 5). While this did not translate into higher root or total dry matter production compared to other genotypes (Table 4), this leaf N% and root nitrate accumulation could be useful for recovery during temporary N limitation during the vegetative stage. This high root nitrate concentration diminished by 65% after three days of nitrate starvation, with no reduction in root dry weight, and these responses were similar to the other genotypes (Table 3). Nitrate in roots and shoots is mainly stored in vacuoles (De Angeli *et al.*, 2006) and contrary to a widely held view, vacuolar nitrate is not a very significant long-term store of N for rapidly growing plants (Clarkson and Hawkesford, 1993).

The root zone nitrate concentration had a highly significant ($P<0.001$) effect on the total N and nitrate concentration measured in the leaves, stems and roots (Table 3). The N concentration in the stems decreased by 76% and in the leaves by 54% in the low N treatment compared to the high N treatment, and corresponded to decreases of 46% and 52% in the dry weight of the stems and leaves respectively (Table 3). The N concentration in roots decreased by 42% and the root dry weight by 26%, in the low N treatment. These results conform to the general observation of a positive relationship between tissue N concentrations and dry matter production (Marschner, 1995).

Starvation of nitrate for 3 days resulted in significant reductions in shoot (leaf & stem) N concentrations, but not is roots (Table 3). This indicates that preference is given to maintaining root N concentrations to ensure continued root growth for nutrient foraging and absorption. This starvation also reduced nitrate concentrations in all tissues and this is understandable given that nitrate is one of the storage forms of nitrogen.

The proportion of the N content in the roots (compared to the total plant N) were in the range of 19-29% in the low N treatment and 11-16% in the high N treatment. This indicates that for young sorghum seedlings, root yields of N can be significant proportions of total plant N yield. Myers (1980) reported similar results for field grown sorghum root N and phosphorus contents

and commented that sorghum roots should not be neglected in drawing up nutrient balance sheets. The range of N concentrations in the roots in the high N treatment was from 2.5 to 3.0%. This range of root N concentrations is similar to the >2% N that resulted in the least germination of seeds of *Striga hermonthica* (Ayongwa *et al.*, 2006), an obligate root hemiparasite that causes serious problems in sorghum production in semi-arid countries (Parker, 2009).

The nitrate concentrations in the stems in the high N treatment (Table 3) were above the value estimated to be toxic in livestock feed (2,100 mg nitrate-N/kg, Stuart 2012). However, it is highly unlikely that in the field such young plants would be grazed or harvested for silage. In addition, the shoot (leaf + stem) nitrate concentration is 1,480 mg nitrate-N/kg; less than the proposed toxic level.

N uptake and utilisation efficiency

As could be expected, N uptake was reduced with reduced nitrate concentrations in the root zone (Table 3). The two genotypes R03128-66 (NAM progeny) and ICSV74 (non-recurrent parent) recorded higher N uptake than R931945-2-2 (Table 6). It is interesting to note that both these two genotypes contain both the NR and GOGAT alleles from ICSV745, but these genotypic differences in N uptake did not result in any genotypic differences in vegetative N use efficiency as measured by the utilisation index (see below).

A number of indices have been proposed to describe various aspects of nitrogen use efficiency (Good *et al.*, 2004; Xu *et al.*, 2012). We chose to use the modified utilisation index (UI) proposed by Siddiqi and Glass (1981) which takes into account both biomass and nutrient concentrations, and is suitable for vegetative growth (see Siddiqi and Glass (1981) for theoretical validity and advantages). Utilization indices were calculated for either the total plant or shoot only (leaf plus stem) (Table 5). There were no significant differences between genotypes in either the total plant N or shoot N utilisation indices (Table 3). The UI were highest in the lowest nitrate treatment and increased in both the high N treatments. This is consistent with the observation that in general, NUE are higher at low N supplies than at high N supplies (Xu et al. 2012). All plants exhibit an increase in NUE under nutrient stress due to, reduced nutrient storage reserves in vacuoles, increased fibre and carbohydrate concentrations, and because a larger proportion of plant biomass is allocated to tissues with low nutrient concentrations (e.g. roots as contrasted to shoots) (Chapin, 1987).

GOGAT gene expression

The expression of only the three GOGAT genes was investigated, the DNA sequences of the three NR genes being too closely related to design discriminating primers (72-81% identity). The expression of the three GOGAT genes (Sobic.009G225700 – NADH-GOGAT, Sobic.003G258800 – NADH-GOGAT, Sobic.002G402700 – Fd-GOGAT) was not significantly influenced by genotype (Table 7). There was however, a significant effect of nitrate concentration in the root zone (Table 7), with the expression of the three GOGAT genes higher in both the high N treatments than in the low N treatment (Table 8A). This result is consistent with increased GOGAT enzyme activity with increased N assimilation (Esposito *et al.*, 2005). Nitrate starvation for 3 days reduced the expression of the three genes (Table 8A), indicating a decrease in N assimilation within a short period of time.

The relative expression of the three GOGAT genes was in the order: Sobic.002G402700> Sobic.003G258800>>>Sobic.009G225700 (Table 8B). The gene Sobic.009G225700 which is under selection was expressed at very low levels in all tissue samples and unaffected by genotype (Tables 5B, 9 & 10). The highest expression of any of the three GOGAT genes occurred for Sobic.002G402700 in both the 10 mM nitrate treatments in the leaves (Table 9) and would indicates its relative importance in N assimilation and metabolism. Physical mapping, sequencing, annotation and candidate gene validation of an NUE metaQTL on wheat chromosome 3B allowed (Quraishi *et al.*, 2011) to propose that a GOGAT gene (homologous to Sobic.003G258800) that is conserved structurally and functionally at orthologous positions in rice, sorghum and maize genomes may contribute to NUE in wheat and other cereals. Gelli et al. (2014) used RNA sequencing to detect common differentially expressed genes in sorghum genotypes with differing sensitivities to low N, grown under different N regimes. The authors found differences in Sobic.003G258800 expression in various genotypes, but these expression levels did not relate to N responsiveness.

Recently a sorghum transcriptome database has been assembled and published (Makita *et al.*, 2015). An examination of this database shows that the highest expression of the three GOGAT genes are in the roots and shoot for Sobic.003G258800, in the leaves for Sobic.002G402700, and in the pistil, endosperm and embryo for Sobic.009G225700. This expression for Sobic.009G225700 (GOGAT under selection) in the reproductive organs would indicate that further experimentation is required to grow the genotypes used in this study to maturity to examine whether there are genotypic differences in yield. The plants grown in this

study were harvested after 30 d, at the 5-6 leaf stage, with the growing point most likely on the verge of differentiation. Floral initiation occurs 30-35 d after emergence (Vanderlip and Reeves, 1972).

Conclusions

Plants respond to nutrient stress with compensatory adjustments such as increased root weight ratio and reduced growth, first in leaf elongation and later in dry weight accumulation (Marschner, 1995). These responses have been reported in a wide range of plant species and the results in the present experiment for nine sorghum genotypes containing varying combinations of allelic variation for a nitrate reductase and a glutamate synthase gene, also conform to these responses. However, despite this, the incorporation of this allelic variation for these two genes into the genome of the elite Australian breeding line did not confer any advantages to this genotype in any of the growth parameters measured. Genotype differences were detected in some of the tissue N and nitrate measurements, but this did not result in any genotypic differences in vegetative nitrogen use efficiency (NUE). The expression of the glutamate synthase gene was very low and unaffected by the treatments, and has recently been reported to be highest in reproductive organs of sorghum. This indicates that further experimentation by growing the genotypes used in this present study to maturity is required, to determine whether there maybe any genotypic differences in yield, NUE or seed protein levels.

CHAPTER 4

CONCLUSIONS AND FUTURE PROSPECTS

4.1 Bioinformatics

A bioinformatic approach identified a nitrate reductase and a NADH-glutamate synthase as being under balancing selection in the genomes of 44 diverse sorghum genotypes (Chapter 2). Genes identified with signatures of balancing selection maintain several alleles in the population conferring a selective heterozygous advantage and likely function optimally under varied conditions according to specificities in development, tissue, or stressors (Massel et al. 2016). Members of the NR and GOGAT gene family have different isoforms with distinct physiological functions, where various genes display cell-specific and organ-specific patterns of expression with differential regulation (Campbell 1999; Suzuki and Knaff, 2005), and as such, the NR and GOGAT genes may play important roles under variable N conditions (Massel et al., 2016). In addition to the NR and GOGAT genes under balancing selection investigated in this thesis, three other genes were also found to be under balancing selection (Chapter 2, Table 2.3). These genes belong to the NPF (NRT/PTR Family = Nitrate transporter 1/Peptide transporter; Leran et al., 2014), and have also been detected by Massel et al. (2016). This provides another set of genes to search for genotypes containing allelic variation for these genes, and phenotype these genotypes for responsiveness to N levels. There are a large number of sorghum genotypes and diversities that have not been investigated, but have been subjected to molecular markers studies (eg. Brenton et al., 2016). When the genomes of such diversity are sequenced, this will provide more opportunities to analyse these genomes to search for genes that are involved in N metabolism and that are also under selection.

4.2 Phenotyping in glasshouse experiment

4.2.1 Growth conditions

The glasshouse experiment was conducted during August in Brisbane Queensland, when night temperatures were sub-optimal for sorghum growth outside the glasshouse (Mean minimum 10.8°C, http://www.bom.gov.au/). In Qld. and northern NSW, sorghum planting time varies from September to January, depending on planting rains and soil temperature early in the season (https://www.daf.qld.gov.au/plants/field-crops-and-pastures/broadacre-field-crops/sorghum/planting-information). The glasshouse experiment was conducted at constant 25°C, no other glasshouse conditions being available at the time. The range of RGRs (relative growth rates, g g^{-1} dry weight) for grain sorghum in this glasshouse experiment was calculated to be 0.05-

0.09 for low N (1 mM N) and 0.08-0.11 for high N (10 mM N). This compares to the range of forage sorghum RGRs reported by Neilson et al. (2015, Figure S3) of 0.05-0.10 for low N (1.5 mM N), 0.12-0.15 for mid N (4.5 mM M) and 0.16-0.19 for high N (20 mM N) (28°C day/18°C night, 700 μmol m^{-2} s^{-1}, during January-March in Adelaide in controlled environment facility). Comparison of these two studies indicates that there may be growth reductions at higher N levels when plants are exposed to continuous 25°C compared to 28°C day/18°C night. However, in a study directly comparing 21°C and 26°C night temperatures (27°C daytime), Manunta and Kirkham (1996) found no differences in sorghum height between plants subjected to either night temperature for 50 days (after transplanting) in a controlled environment room. Even after 74 days, there were still no differences between 21°C and 26°C night time treatments in the fresh and dry weights of stems and leaves.

4.2.2 Gene expression

Whilst it was possible to design gene specific primers for the three GOGAT genes in the sorghum genomes to examine gene expression, this was not possible for the three NR genes due to very homologous DNA sequences and the use of simple traditional primer design (Chapter 3.0). Holland et al. (1991) described a procedure to produce gene specific probes that increase the specificity of the qPCR reaction, and this technique is one of the most popular with commercial kits now being available (eg. Taqman®, Thermo Fisher Scientific). This technique is also able to distinguish single nucleotide polymorphisms and splice variants.

4.3 Future phenotyping experiments

None of the tested genotypes exceeded growth or NUE of elite parent R931945-2-2, indicating that the allelic combinations did not confer an advantage during early vegetative growth. In addition, the gene expression of the GOGAT under balancing selection and for which allelic variation was tested, is highest in reproductive tissues. Nitrogen dynamics during this reproductive stage are dominated by recycling and remobilisation processes, and are very different from the N dynamics during the vegetative stage, which are dominated by absorption, assimilation and metabolism (Hirel et al., 2001, 2007). Consequently, the next steps for ascertaining potential effects on NUE include growing plants to maturity. This can be achieved by growing the sorghum genotypes to maturity in a glasshouse, ensuring appropriate temperature regimes (see previous section), such as 28°C day/22°C night used by Kim et al (2010) to grow sorghum to maturity. However, while such glasshouse experiments will give an indication of yield and grain characteristics, more detailed field experiments are required to obtain real world results. Field trials

will also allow increased replication and the potential for including other genotypes of interest (see section "4.1 Bioinformatics" above). Field trials examining yield in sorghum NAM populations are well established in Qld. and NSW (Jordan et al., 2011). In addition, field trials examining N dynamics in sorghum have been conducted in Qld. (eg. van Oosterom et al., 2010 a, b) and hence field locations and experimental logistics are also well established. Nitrogen dynamics have been described in more detail using N^{15} labelling in maize (Gallais et al., 2006) and this technique can easily be adapted to sorghum grown under both glasshouse and field conditions.

When growing plants to maturity, consideration needs to be given to the maturity dates of the genotypes being tested. This is important to ensure that genotypic comparisons are made at similar growth stages. This has been emphasized by Sadras and Lemaire (2014) with respect to the N nutrition of crops. The recurrent parent of the NAM progenies tested in this thesis is R31945-2-2, with the progenies containing 79-85% of the R31945-2-2 genome (Chapter 2.0, Table 2.4). Consequently, it can be assumed that both R31945-2-2 and the NAM progenies will have similar intervals between similar growth stages of successive leaves (phyllochron) and also during reproductive development and maturity dates. However, special attention will need to be given to the plant tissue sampling times of the other two genotypes, ISCV745 and Macia. ISCV745 has a 10 day difference in maturity relative to R31945-2-2, and Macia, a four day difference in maturity relative to R31945-2-2 (Jordan et al., 2011, Table 1). Data from 21 field trial locations indicated R31945-2-2 takes approximately 59.5 days to flower, ISCV745 58 days, and Macia 58.5 days (Jordan et al., 2011, Table 1).

This thesis has shown how the combined use of forward genetics, genotype diversity and phenotyping, can take advantage of the rapidly expanding genomic databases to enable a systematic approach for investigating and developing N efficient crops.

REFERENCES

Asher CJ, Cowie AM. 1974. Grain sorghum — high yield, satisfactory protein content or both? Proceedings of the Agronomy Society of New Zealand **4**, 79-82.

Ayongwa GC, Stomph TJ, Emechebe AM, Kuyper TW. 2006. Root nitrogen concentration of sorghum above 2% produces least Striga hermonthica seed stimulation. Annals of Applied Biology **149**, 255-262.

Barnaud A, Trigueros G, McKey D, Joly HI. 2008. High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? Heredity **101**, 445-452.

Bekele WA, Wieckhorst S, Friedt W, Snowdon RJ. 2013. High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. Plant Biotechnol J **11**, 1112-1125.

Bloom AJ, Burger M, Kimball BA, Pinter PJ. 2014. Nitrate assimilation is inhibited by elevated CO_2 in field-grown wheat. Nature Climate Change **4**, 477–480.

Bloom AJ, Rubio-Asensio JS, Randall L, Rachmilevitch S, Cousins AB, Carlisle EA. 2012. CO2 enrichment inhibits shoot nitrate assimilation in C-3 but not C-4 plants and slows growth under nitrate in C-3 plants. Ecology **93**, 355-367.

Brenton ZW, Cooper EA, Myers MT, Boyles RE, Shakoor N, Zielinski KJ, Rauh BL, Bridges WC, Morris GP, Kresovich S. 2016. A Genomic Resource for the Development, Improvement, and Exploitation of Sorghum for Bioenergy. Genetics **204**, 21-33.

Brown RH. 1978. Difference in N use efficiency in C3 and C4 plants and its implications in adaptation and evolution. Crop Science **18**, 93-98.

Campbell WH. 1999. Nitrate reductase structure, function and regulation: Bridging the gap between biochemistry and physiology. Annual Review of Plant Physiology and Plant Molecular Biology **50**, 277-+.

Chapin F. 1987. Adaptations and physiological responses of wild plants to nutrient stress. Genetic aspects of plant mineral nutrition, 15-25.

Clarkson DT, Hawkesford MJ. 1993. Molecular biological approaches to plant nutrition. Plant and Soil **155**, 21-31.

Cowie AM. 1973. Effects of nitrogen supply on grain yield and protein content in hybrid grain sorghum., University of Queensland.

De Angeli A, Monachello D, Ephritikhine G, Frachisse JM, Thomine S, Gambale F, Barbier-Brygoo H. 2006. The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles. Nature **442**, 939-942.

Delph LF, Kelly JK. 2014. On the importance of balancing selection in plants. New Phytologist **201**, 45-56.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research **36**, W465-W469.

Dillon SL, Shapter FM, Henry RJ, Cordeiro G, Izquierdo L, Lee LS. 2007. Domestication to crop improvement: Genetic resources for Sorghum and saccharum (Andropogoneae). Annals of Botany **100**, 975-989.

Dodd IC, Munns R, Passioura JB. 2002. Does shoot water status limit leaf expansion of nitrogen-deprived barley? Journal of Experimental Botany **53**, 1765-1770.

Erisman JW, Sutton MA, Galloway J, Klimont Z, Winiwarter W. 2008. How a century of ammonia synthesis changed the world. Nature Geoscience **1**, 636-639.

Esposito S, Guerriero G, Vona V, Rigano VD, Carfagna S, Rigano C. 2005. Glutamate synthase activities and protein changes in relation to nitrogen nutrition in barley: the dependence on different plastidic glucose-6P dehydrogenase isoforms. Journal of Experimental Botany **56**, 55-64.

Fischer RA, Byerlee D, Edmeades GO. 2014. *Crop yields and global food security. Will yield increase continue to feed the world?*: ACIAR Monograph No. 158. Australian Centre for International Agricultural Research: Canberra. xxii.

Forno DA. 1977. The mineral nutrition of cassava (*Manihot esculenta* Crantz) with particular reference to nitrogen., University of Queensland.

Gallais A, Coque M, Quillere I, Prioul JL, Hirel B. 2006. Modelling postsilking nitrogen fluxes in maize (Zea mays) using N-15-labelling field experiments. New Phytologist **172**, 696-707.

Gardner JC, Maranville JW, Paparozzi ET. 1994. Nitrogen use efficiency among diverse sorghum cultivars. Crop Science **34**, 728-733.

Gelli M, Mitchell SE, Liu K, Clemente TE, Weeks DP, Zhang C, Holding DR, Dweikat IM. 2016. Mapping QTLs and association of differentially expressed gene transcripts for multiple agronomic traits under different nitrogen levels in sorghum. Bmc Plant Biology **16**, 18.

Gelli MD, Y.; Konda AR; Zhang C; Holding D; Dweikat I. 2014. Identification of differentially expressed genes between sorghum genotypes with contrasting nitrogen stress tolerance by genome-wide transcriptional profiling. BMC Genomics **15**, 1-16.

Ghannoum O, Evans JR, von Caemmerer S. 2011. Nitrogen and water use efficiency of C_4 plants. C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms, 129–146.

Good AG, Shrawat AK, Muench DG. 2004. Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science **9**, 597-605.

Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research **36**, 3420-3435.

Gruber N, Galloway JN. 2008. An Earth-system perspective of the global nitrogen cycle. Nature **451**, 293-296.

Grundon NE, DG;Takkar, PN; Asher, CJ;Clark, RB. 1987. *Nutritional disorders of grain sorghum*: Australian Center for International Agricultural Research.

Hallock DL, Brown RH, Blaser RE. 1965. Relative Yield and Composition of Ky. 31 Fescue and Coastal Bermudagrass at Four Nitrogen Levels. Agronomy Journal **57**, 539-542.

Harlan JR, Dewet JMJ. 1972. Simplified classification of cultivated sorghum. Crop Science **12**, 172-&.

He ZL, Zhang HK, Gao SH, Lercher MJ, Chen WH, Hu SN. 2016. Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Research **44**, W236-W241.

Hibberd DE, Hall BD. 1990. The responses of maize and grain-sorghum hybrids to nitrogen-fertilizer in south east Queensland. Australian Journal of Experimental Agriculture **30**, 825-831.

Hirel B, Bertin P, Quillere I, Bourdoncle W, Attagnant C, Dellay C, Gouy A, Cadiou S, Retailliau C, Falque M, Gallais A. 2001. Towards a better understanding of the genetic and physiological basis for nitrogen use efficiency in maize. Plant Physiology **125**, 1258-1270.

Hirel B, Le Gouis J, Ney B, Gallais A. 2007. The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. Journal of Experimental Botany **58**, 2369-2387.

Holland PM, Abramson RD, Watson R, Gelfand DH. 1991. Detection of specific polymerase chain-reaction product by utilizing the 5'- 3' exonuclease activity of Thermus-aquaticus DNA-polymerase. Proceedings of the National Academy of Sciences of the United States of America **88**, 7276-7280.

Jordan DR, Hunt CH, Cruickshank AW, Borrell AK, Henzell RG. 2012. The relationship between the stay-green trait and grain yield in elite sorghum hybrids grown in a range of environments. Crop Science **52**, 1153-1161.

Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Henzell RG. 2011. Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. Crop Science **51**, 1444-1457.

Jordan DR, Tao YZ, Godwin ID, Henzell RG, Cooper M, McIntyre CL. 1998. Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. Euphytica **102**, 1-7.

Kamoshita A, Fukai S, Muchow RC, Cooper M. 1998a. Genotypic variation for grain yield and grain nitrogen concentration among sorghum hybrids under different levels of nitrogen fertiliser and water supply. Australian Journal of Agricultural Research **49**, 737-747.

Kamoshita A, Fukai S, Muchow RC, Cooper M. 1998b. Sorghum hybrid differences in grain yield and nitrogen concentration under low soil nitrogen availability - II. Hybrids with contrasting phenology. Australian Journal of Agricultural Research **49**, 1277-1286.

Kamoshita A, Fukai S, Muchow RC, Cooper M. 1998c. Sorghum hybrid differences in grain yield and nitrogen concentration under low soil nitrogen availability. I. Hybrids with similar phenology. Australian Journal of Agricultural Research **49**, 1267-1276.

Kamoshita A, Fukai S, Muchow RC, Cooper M. 1999. Grain nitrogen concentration differences among three sorghum hybrids with similar grain yield. Australian Journal of Agricultural Research **50**, 137-145.

Kellogg EA. 2013. Phylogenetic relationships of Saccharinae and Sorghinae. Genomics of the Saccharinae **11**, 3-21.

Kiirats O, Kramer DM, Edwards GE. 2010. Co-regulation of dark and light reactions in three biochemical subtypes of C-4 species. Photosynthesis Research **105**, 89-99.

Kim HK, van Oosterom E, Dingkuhn M, Luquet D, Hammer G. 2010. Regulation of tillering in sorghum: environmental effects. Annals of Botany **106**, 57-67.

Kopriva S. 2011. Nitrogen and Sulfur Metabolism in C_4 Plants. C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms, 109–128.

Lawrence CJ, Walbot V. 2007. Translational genomics for bioenergy production from fuelstock grasses: maize as the model species. Plant Cell **19**, 2091-2094.

Lea PJ, Blackwell RD, Chen FL, Hecht U. 1990. Enzymes of ammonia assimilation. Methods in Plant Biochemistry, Enzymes of primary metabolism **3**, 257-276.

Lemaire G, Charrier X, Hebert Y. 1996. Nitrogen uptake capacities of maize and sorghum crops in different nitrogen and water supply conditions. Agronomie **16**, 231-246.

Leran S, Varala K, Boyer JC, Chiurazzi M, Crawford N, Daniel-Vedele F, David L, Dickstein R, Fernandez E, Forde B, Gassmann W, Geiger D, Gojon A, Gong JM, Halkier BA, Harris JM, Hedrich R, Limami AM, Rentsch D, Seo M, Tsay YF, Zhang MY, Coruzzi G, Lacombe B. 2014. A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. Trends in Plant Science **19**, 5-9.

Luo H, Zhao W, Wang Y, Xia Y, Wu X, Zhang L, Tang B, Zhu J, Fang L, Du Z, Bekele WA, Tai S, Jordan DR, Godwin ID, Snowdon RJ, Mace ES, Luo J, Jing H-C. 2016. SorGSD: a sorghum genome SNP database. Biotechnology for Biofuels **9**, 6.

Mace ES, Tai SS, Gilding EK, Li YH, Prentis PJ, Bian LL, Campbell BC, Hu WS, Innes DJ, Han XL, Cruickshank A, Dai CM, Frere C, Zhang HK, Hunt CH, Wang XY, Shatte T, Wang M, Su Z, Li J, Lin XZ, Godwin ID, Jordan DR, Wang J. 2013. Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. Nature Communications **4**, 9.

Makita Y, Shimada S, Kawashima M, Kondou-Kuriyama T, Toyoda T, Matsui M. 2015. MOROKOSHI: Transcriptome Database in Sorghum bicolor. Plant and Cell Physiology **56**, 8.

Manunta P, Kirkham MB. 1996. Respiration and growth of sorghum and sunflower under predicted increased night temperatures. Journal of Agronomy and Crop Science **176**, 267-274.

Maranville JW, Clark RB, Ross WM. 1980. Nitrogen efficiency in grain-sorghum. Journal of Plant Nutrition **2**, 577-589.

Maranville JW, Madhavan S. 2002. Physiological adaptations for nitrogen use efficiency in sorghum. Plant and Soil **245**, 25-34.

Maranville JW, Pandey RK, Sirifi S. 2002. Comparison of nitrogen use efficiency of a newly developed sorghum hybrid and two improved cultivars in the Sahel of West Africa. Communications in Soil Science and Plant Analysis **33**, 1519-1536.

Marschner H. 1995. *Mineral nutrition of higher plants*: Academic Press.

Martin F, Suzuki A, Hirel B. 1982. A new high-performance liquid-chromatography assay for glutamine-synthetase and glutamate synthase in plant-tissues. Analytical Biochemistry **125**, 24-29.

Massel K, Campbell BC, Mace ES, Tai S, Tao Y, Worland BG, Jordan DR, Botella JR, Godwin ID. 2016. Whole genome sequencing reveals potential new targets for improving nitrogen uptake and utilization in sorghum bicolor. Frontiers in Plant Science **7**.

Miranda KM, Espey MG, Wink DA. 2001. A rapid, simple spectrophotometric method for simultaneous detection of nitrate and nitrite. Nitric Oxide-Biology and Chemistry **5**, 62-71.

Muchow RC. 1988. Effect of nitrogen supply on the comparative productivity of maize and sorghum in a semi-arid tropical environment. 1. Leaf growth and leaf nitrogen. Field Crops Research **18**, 1-16.

Muraya MM. 2014. Sorghum Genetic Diversity. Genetics, Genomics and Breeding of Sorghum, 114-140.

Myers RJK. 1980. The root-system of a grain-sorghum crop. Field Crops Research **3**, 53-64.

Nei M. 1987. *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

Neilson EH, Edwards AM, Blomstedt CK, Berger B, Moller BL, Gleadow RM. 2015. Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C-4 cereal crop plant to nitrogen and water deficiency over time. Journal of Experimental Botany **66**, 1817-1832.

Palmer SJ, Berridge DM, McDonald AJS, Davies WJ. 1996. Control of leaf expansion in sunflower (Helianthus annuus L) by nitrogen nutrition. Journal of Experimental Botany **47**, 359-368.

Pang JY, Palta JA, Rebetzke GJ, Milroy SP. 2014. Wheat genotypes with high early vigour accumulate more nitrogen and have higher photosynthetic nitrogen use efficiency during early growth. Functional Plant Biology **41**, 215-222.

Parker C. 2009. Observations on the current status of Orobanche and Striga problems worldwide. Pest Management Science **65**, 453-459.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang HB, Wang XY, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang LF, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KFX, Messing J, Rokhsar DS. 2009. The Sorghum bicolor genome and the diversification of grasses. Nature **457**, 551-556.

Pearcy RW, Ehleringer J. 1984. Comparative ecophysiology of C-3 and C-4 plants. Plant Cell and Environment **7**, 1-13.

Quraishi UM, Abrouk M, Murat F, Pont C, Foucrier S, Desmaizieres G, Confolent C, Riviere N, Charmet G, Paux E, Murigneux A, Guerreiro L, Lafarge S, Gouis J, Feuillet C, Salse J. 2011. Cross-genome map based dissection of a nitrogen use efficiency ortho-metaQTL in bread wheat unravels concerted cereal genome evolution. Plant Journal **65**, 745-756.

Reuter DR, JB. 1997. *Plant analysis: an interpretation manual*: CSIRO Publishing.

Reynolds HL, Dantonio C. 1996. The ecological significance of plasticity in root weight ratio in response to nitrogen: Opinion. Plant and Soil **185**, 75-97.

Robinson N, Brackin R, Vinall K, Soper F, Holst J, Gamage H, Paungfoo-Lonhienne C, Rennenberg H, Lakshmanan P, Schmidt S. 2011. Nitrate paradigm does not hold up for sugarcane. PLoS ONE **6**, 9.

Robinson N SSaLP. 2015. Genetic improvement of nitrogen use efficiency in sugarcane. A review of nitrogen use efficiency in sugarcane, 125-156.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research **37**, 12.

Sadras VO, Lemaire G. 2014. Quantifying crop nitrogen status for comparisons of agronomic practices and genotypes. Field Crops Research **164**, 54-64.

Sage RF, Pearcy RW, Seemann JR. 1987. The nitrogen use efficiency of C-3 and C-4 plants. 3. Leaf nitrogen effects on the activity of carboxylating enzymes in Chenopodium-album (L) and Amaranthus-retroflexus (L). Plant Physiology **85**, 355-359.

Siddiqi MY, Glass ADM. 1981. Utilization index - a modified approach to the estimation and comparison of nutrient utilization efficiency in plants. Journal of Plant Nutrition **4**, 289-302.

Singh HP, Lohithaswa HC. 2006. Sorghum. Genome Mapping and Molecular Breeding in Plants; Cereals and Millets.

Singh V, van Oosterom EJ, Jordan DR, Hunt CH, Hammer GL. 2011. Genetic variability and control of nodal root angle in sorghum. Crop Science **51**, 2011-2020.

Skendzic EMC, J.T; and Cerros-Tlatilpa, R. 2013. Phylogenetics of Andropogoneae (Poaceae: Panicoideae) Based on Nuclear Ribosomal Internal Transcribed Spacer and Chloroplast trnL–F Sequences. Aliso: A Journal of Systematic and Evolutionary Botany **23**.

Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Research **33**, W465-W467.

Stuart NS. 2012. Toxicology and associated agronomy of forage sorghum. PhD, University of Queensland, 284.

Suzuki A, Audet C, Oaks A. 1987. Influence of light on the ferredoxin-dependent glutamate synthase in maize leaves. Plant Physiology **84**, 578-581.

Suzuki A, Jacquot JP, Martin F, Gadal P. 1982. Light-dependent activity of glutamate synthase *in vitro*. Biochemical and Biophysical Research Communications **106**, 65-71.

Suzuki A, Knaff DB. 2005. Glutamate synthase: structural, mechanistic and regulatory properties, and role in the amino acid metabolism. Photosynthesis Research **83**, 191-217.

Tajima F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**, 585-595.

Tamura W, Kojima S, Toyokawa A, Watanabe H, Tabuchi-Kobayashi M, Hayakawa T, Yamaya T. 2011. Disruption of a novel NADH-glutamate synthase2 gene caused marked reduction in spikelet number of rice. Frontiers in Plant Science **2**, 9.

Tian DC, Araki H, Stahl E, Bergelson J, Kreitman M. 2002. Signature of balancing selection in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **99**, 11525-11530.

van Oosterom EJ, Borrell AK, Chapman SC, Broad IJ, Hammer GL. 2010a. Functional dynamics of the nitrogen balance of sorghum: I. N demand of vegetative plant parts. Field Crops Research **115**, 19-28.

van Oosterom EJ, Chapman SC, Borrell AK, Broad IJ, Hammer GL. 2010b. Functional dynamics of the nitrogen balance of sorghum. II. Grain filling period. Field Crops Research **115**, 29-38.

Vanderlip RL, Reeves HE. 1972. Growth stages of sorghum Sorghum bicolor, (L) Moench. Agronomy Journal **64**, 13-+.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology **3**, 12.

Warncke DD, Barber SA. 1974. Nitrate uptake effectiveness of 4 plant species. Journal of Environmental Quality **3**, 28-30.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology **7**, 256-276.

Williams RF. 1948. The effects of phosphorus supply on the rates of intake of phosphorus and nitrogen and upon certain aspects of phosphorus metabolism in gramineous plants. Australian Journal of Scientific Research Series B-Biological Sciences **1**, 333-361.

Wray JL FR. 1990. Nitrate reductase and nitrite reductase. Methods in Plant Biochemistry **3**, 241–256.

Wright SI, Gaut BS. 2005. Molecular population genetics and the search for adaptive evolution in plants. Molecular Biology and Evolution **22**, 506-519.

Xu GH, Fan XR, Miller AJ. 2012. Plant Nitrogen Assimilation and Use Efficiency. Annual Review of Plant Biology, Vol 63 **63**, 153-182.

Yamaya T, Obara M, Nakajima H, Sasaki S, Hayakawa T, Sato T. 2002. Genetic manipulation and quantitative-trait loci mapping for nitrogen recycling in rice. Journal of Experimental Botany **53**, 917-925.

Yang Z, Hammer, G., van Oosterom, E., Rochais, D., Deifel, K. . 2010. Effects of pot size on growth of maize and sorghum plants Proceedings of the 1st Australian summer grains conference.

Youngquist JB, Bramelcox P, Maranville JW. 1992. Evaluation of alternative screening criteria for selecting nitrogen-use efficient genotypes in sorghum. Crop Science **32**, 1310-1313.

APPENDIX

Table 1S Sequences of primers used in qPCR analyses.

Primer identification	sequence 5'→3'
Sobic.009G225700-5899F	TTGGGACAGCCATCAGACATGG
Sobic.009G225700-6015R	TTCGGAATATCCTCGGCCACTGAG
Sobic.003G258800-3599F	TTCGTGGTCGAGCAGTTCTGC
Sobic.003G258800-3679R	CTGCACCAAGCAAGCAAGCAAC
Sobic.002G402700-1710F	AGAAGCTGACGCTGCTGTGC
Sobic.002G402700-1795R	TGGCAGGACGAGTTGGTTCAGG
Sobic.009G174900-F	CGACCAGCAACAAACCCAAG
Sobic.009G174900-R	CCCTGAGATTGCCCACATGT
Sobic.002G339600-F	CCCAAGTACTCCAAGGCTCG
Sobic.002G339600-R	ATGTTGTCACCCTCGAACCC